Université de Montréal

Consumption of black tea and coffee and the risk of lung cancer

Par
Romain Pasquet

Département de médecine sociale et préventive
Faculté de Médecine

Mémoire présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade
M.Sc. en Santé Communautaire

Décembre 2012

Université de Montréal
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé:

Consumption of black tea and coffee and the risk of lung cancer

Présenté par :
Romain Pasquet

a été évaluée par un jury composé des personnes suivantes :

Helen Trottier, président-rapporteur
Anita Koushik, directeur de recherche
Igor Karp, co-directeur
Marie-Claude Rousseau, membre du jury

Résumé

Objectif: Étudier l'association entre la consommation de café, la consommation de thé noir et le risque de cancer du poumon.

Méthodologie: Cette recherche utilise les données obtenues lors d'une étude cas-témoin effectuée à Montréal avec des résidents canadiens âgés entre 35 et 75 ans recrutés entre 1996 et 2001. Les cas étaient des individus atteints du cancer et diagnostiqués entre 1996 et 1997 dans l'un des 18 hôpitaux de la région du Grand-Montréal. Les contrôles on été sélectionnés à partir de la liste électorale et appariés selon la fréquence de distribution des cas par groupe d'âge de 5 ans, par sexe et par district électoral. Au total, 1130 cas et 1484 contrôles ont été inclus dans cette étude. Les rapports de cote (RC) et les intervalles de confiance de 95% (CI) des associations entre la consommation de thé noir, de café et le cancer du poumon ont été estimés à l'aide de la régression logistique non conditionnelle. Quatre aspects de la consommation ont été analysés dans des modèles multivariés distincts: la fréquence de consommation, la consommation journalière moyenne, la durée de consommation et la consommation cumulative. Les covariables potentielles considérées incluaient : l'âge, le sexe, l'historique de tabagisme, le statut du répondant, l'ethnicité, la consommation d'alcool, la consommation de fruit et de légume, l'apport énergétique journalier, l'exposition a des agents professionnelle et les variables socio-économiques . Des analyses secondaires par le sexe, le tabagisme et le type histologique de cancer on été effectuées.

Résultats : Aucune association statistiquement significative n'a été observée entre la consommation de thé noir et le cancer du poumon. Nos résultats suggèrent qu'une consommation de $\geq$ 50 ans était associée avec une augmentation du risque d'adénocarcinome comparée à aucune consommation. Nous avons observé une association inverse statistiquement significative entre la consommation occasionnelle de café et le cancer du poumon (RC : 0.32, 95%CI : 0.17-0.59). La durée de consommation de café n'était pas associée avec le cancer du poumon.

Conclusion : Nos résultats suggèrent que la consommation à long terme de thé noir pourrait augmenter le risque d'adénocarcinome. D'un autre côté, la consommation occasionnelle de café pourrait réduire le risque de cancer du poumon.

Mots-clés : Épidémiologie, cas-contrôle, poumon, cancer, thé, thé noir, café

Abstract

Objective: To investigate the associations between the consumption of black tea and coffee and lung cancer.

Methods: This research was conducted using data from a Montreal lung cancer case-control study that included Canadian residents aged 35 to 75 years old recruited between 1996 and 2001. Cases were individuals diagnosed with lung cancer between 1996 and 1997 from one of 18 Montreal-area hospitals. Controls were randomly selected from the electoral list and frequency matched to the distribution of the cases by 5 year age groups, sex and electoral district. In total, 1130 cases and 1484 controls were included in this analysis. The odds ratios (OR) and 95% confidence intervals (CI) for the association between black tea and coffee consumption and lung cancer were estimated using unconditional logistic regression. Four aspects of the consumption were analyzed in separate multivariate models: the frequency of consumption, average daily amount of consumption, duration of consumption and cumulative consumption. Potential covariates included: age, sex, smoking, respondent status, ethnicity, alcohol intake, fruit and vegetable intake, energy intake, exposure to occupational agents and socioeconomic variables. Analyses by sex, smoking level and tumor histological type were also conducted.

Results: No statistically significant association was observed between the consumption of black tea and lung cancer. There were indications that consuming black tea for 50 years or more was associated with an increase in the risk of adenocarcinoma relative to no consumption. A significant inverse association between the consumption of coffee and lung cancer was observed for occasional coffee drinkers when compared to never drinkers with an estimated OR (95%CI) of 0.32 (0.17 - 0.59). The duration of coffee consumption was not statistically significantly associated with lung cancer.

Conclusion: Our results suggest that long-term consumption of black tea may increase the risk of adenocarcinoma. On the other hand, the occasional consumption of coffee may reduce the risk of lung cancer.

Keywords : epidemiology, case-control, lung, cancer, tea, black tea, coffee

Table of contents

Table list

Figure list

Abbreviations list

NSLC	Non small cell lung cancer
SCLC	Small cell lung cancer
SMR	Standardized mortality ratio
EGCG	Epigallocatechin gallate
IARC	International Agency for Research on Cancer
PKA	Protein kinase A
ERK1/2	Extracellular signal-regulated protein kinase 1 and 2
CREB	cAMP response binding element
RR	Risk ratio
95% CI	95% confidence interval
OR	Odd ratio
HR	Hazard ratio
FFQ	food frequency questionnaire
MLCS	Montreal lung cancer case-control study
Q	Question
CSI	Comprehensive smoking index
LR	Likehood ratio
SAS	Statistical Analysis Software
CC	Case control study

Acknowledgements

I would first and foremost like to express my gratitude to my supervisor, Anita Koushik, and my co-supervisor, Igor Karp. Without their constant support, guidance and infinite patience, I would never have gained the knowledge and skills necessary to see this project to completion.

I am grateful to Jack Siemiatycki, Lesley Richardon and Javier Pintos for spending so much of their free (and busy) time to discuss the various issues I encountered while completing my thesis.

I am also grateful to Sally Campbell for the help with the literature review. I do not know how I would have been able to complete this section without her.

I would like to thank Éric Vallières who by sharing the same office as me was pestered so often with my questions. I will pay back for all those coffees one day I promise.

I would also like to thank Marie-Claude, Alassane, Julie, Jérome, Marie-Élise, Aude and everybody else from the Environmental Epidemiology and Population Health Research Group for providing such a friendly, interesting and fun environment to learn and work in. I enjoyed every moments I spent working and relaxing with all of them.

Last, but not least, I am thankful to Mengting, my family and my friends for supporting me during this past year even though I never returned their calls. I hope they will not hold any grudge against me.

1.0 Introduction

Cancer is one of the leading concerns in public health, it was responsible for 7.6 million deaths worldwide in 2008, representing 13% of declared deaths that year (1). Of all types of cancer, lung cancer is the most diagnosed (2, 3) and caused 1.38 million or 18% of all cancer deaths in 2008 (1). In Canada in 2011, the estimated age-adjusted incidence rate for lung cancer was 65 per 100 000 person-years for males and 51 per 100 000 person-years for females, representing 25 400 cases (13 200 males, 12 200 females). The same year, an estimated 20 600 Canadians (11 300 males and 9 300 females) died from lung cancer (4). The age-adjusted incidence rate of lung cancer in Canada has been decreasing in males from 95.1 per 100 000 person-years during the late 80's to what we see today however the same is not true for females where the estimated rate has increased from the 35 per 100 000 person-years observed in the late 80's (4). Of all provinces, Quebec has the highest rate of lung cancer with an estimated age adjusted incidence rate of 92 per 100 000 person-years in males and 67 per 100 000 person-years in females in 2011 (4).

The principal known risk factor of lung cancer is smoking which is believed to be responsible for over 87% of all lung cancer cases (5). However, lung cancer also occurs in non-smokers and secondary smoke cannot explain all observed cases in this subpopulation (2, 3, 6). Thus, a number of genetic, environmental and occupational factors (3) have been studied and are believed to play a role in the development of lung cancer. Dietary factors have been of interest in cancer epidemiology (7-10, 11) and while the role of diet in lung cancer is still controversial, their potential role cannot be overlooked (10). Tea and coffee are two beverages consumed worldwide by billions of individuals (12). Both are composed of hundred of chemicals, some of which are believed to play a role in the development of lung cancer (12, 13). While the consumption of both tea and coffee have been the subject of epidemiological studies, the evidence for an association with lung cancer has been inconsistent, due in part to the various methodologies used, the lack of consideration of long-term consumption and potential residual confounding by smoking (14, 15). Thus, to

better understand the role of those two beverages in the development of lung cancer further analysis is required.

2.0 Literature review

2.1 Lung cancer

2.1.1 Descriptive epidemiology

Lung cancer is defined as a "*cancer that forms in tissues of the lung, usually in the cells lining air passages*" (16). It develops into two main subtypes; Non small cell lung cancer (NSLC) and small cell lung cancer (SCLC), which represent around 80% and 20% of all lung cancer cases respectively (5, 17). Four types of NSLC (adenocarcinomas, squamous cell carcinomas, large cell carcinomas and bronchioalveolar carcinomas), and two types of SCLC (small cell carcinomas and combined small cell carcinomas) account for most cases (5, 17). During the 70s the most common type of lung cancer was squamous cell carcinomas but today adenocarcinoma has become the most prevalent form in many countries including Japan and the United-States (3). It has been proposed that this shift is the result of a change in inhalation habits caused by the introduction of low-yield and filtered cigarettes (3). Both subtypes have reported 5-year survival rates of 5% to 15% (6, 18). Surgery is the most effective and successful way to treat non-metastatic NSLC however in 70% of the cases the cancer is discovered too late and must be treated using the less effective chemotherapy and radiotherapy treatments (6). SCLC is a more aggressive cancer which is harder to operate but responds better to chemotherapy and radiotherapy (18). However, while chemotherapy treatments have been shown to be very effective, particularly in limited states (i.e. only present in one lung and its nearby tissues or lymph nodes) in SCLC patients, relapse and death is still common (18).

2.1.2 Risk factors

The aetiology of lung cancer has been extensively studied and some risk factors are known and include dietary factors, socioeconomic factors, genetic factors, occupational factors and smoking.

2.1.2.1 Smoking

Smoking is the principal known risk factor of lung cancer. Cigarette smoke is composed of over 80 known carcinogens and could be responsible for more than 87% of lung cancer cases (5). The higher rate of lung cancer in males and the gender specific change in incidence rate observed in the past decades are both believed to be related to changes in smoking habits best described as a general decrease in the amount of male smokers and a general increase in female smokers (3, 4). An increase in either the duration of smoking or in the amount smoked has been positively associated with an increase in the risk of lung cancer (2). While a lower risk of lung cancer has been observed for individuals smoking filtered cigarettes when compared to individuals smoking a similar amount of un-filtered cigarettes, it has also been observed that smokers tend to smoke more filtered cigarettes to make up for the reduction in tar (2, 3). Smoking can also affect non-smokers through passive smoking which has been shown to be positively associated with lung cancer and to be responsible for around 1.6% of lung cancer cases (2, 3, 6).

2.1.2.2 Socioeconomic factors

In males of industrial countries, a higher incidence of lung cancer has been reported in lower socioeconomic (income, education, occupation, social class) groups when compared to higher socioeconomic groups, however in females the evidence is inconsistent (3). It has been proposed that a higher rate of smoking among individuals from lower socioeconomic background could explain a good proportion of the observed difference in cancer rate between individuals from lower and higher socioeconomic background.(19).

Lower incidence rates of lung cancer have been observed in Asians and Hispanics when compared to Caucasians, while higher rates have been observed in Blacks (3, 20). However, the evidence for a genetic or biologic cause is weak and socioeconomic factors are more likely to explain the variation observed between the various ethnicities (20).

2.1.2.3 Genetic factors

Higher risks of lung cancer among family members of lung cancer patients have led to the discovery of many genetic factors positively associated with lung cancer. Polymorphisms in the cytochrome P-450 supergene family of enzymes, which are known to metabolize many carcinogens, have been associated with lung cancer (3). Similarly, polymorphisms in other genes such as the microsomal epoxide hydrolase, NAD(P)H quinone oxydoreductase 1, and in genes associated with methyl metabolism, DNA repair and cell cycle control have also been associated with lung cancer (3).

.

2.1.2.4 Occupational factors

Exposure to asbestos is the principal known occupational risk factor for lung cancer. It has been estimated that exposure to asbestos accounts for up to 7% of all lung cancer cases, with an estimated death toll of 100 000 to 140 000 per year (21). A meta-analysis (22) of the risk of lung cancer in 69 asbestos-exposed occupational cohorts reported a meta-SMR of 163 in cohorts that considered latency, and a meta-SMR of 148 in cohorts where latency was not considered. Exposure to asbestos can occur through the use of asbestos products but is most often associated with asbestos-related jobs (23). Today, maintenance and construction workers are the occupational groups most often exposed to asbestos (3). Between the early 40's and the late 70's, 27 million workers were estimated to have been exposed to asbestos fibers in the United-State alone (21), with an estimated 1.3 million workers still at risk of exposure today (23). While exposure to asbestos is still a common

occurrence in developing countries, steps have been taken in developed countries to reduce exposure (23), but with a latency period between exposure to asbestos and the development of lung cancer of 30 to 40 years, more asbestos-caused lung cancer cases can be expected to appear in the near future (21). Many other occupational agents have been found to be positively associated with lung cancer including: radon, arsenic, chloromethyl ethers, silica, chromium compounds, nickel and polycyclic aromatic hydrocarbons (3).

2.1.2.5 Dietary factors

Various dietary factors have been reported to be associated with lung cancer. A positive association between fat intake and lung cancer is still debated with both positive (8, 9) and null associations (7) having been reported in the literature. Higher alcohol intake has been positively associated with lung cancer (3, 24), but wine was also reported to reduce the risk of lung cancer (24), and confounding by smoking is difficult to rule out (3). Higher consumption of fruits and vegetables has been inversely associated with lung cancer, as has higher intake of carotenoids, folates, isothiocyanates and selenium (3). While inconsistent results have been observed, tea and coffee have also been proposed to play a role in the development of lung cancer (14, 15).

2.2 Tea

Tea is considered one of the oldest beverages in the world. Legend places its discovery to 2700 B.C in China where the emperor Shen Nung tasted the beverage following the fall of a tea leaf in his cup of hot water (25). Tea quickly travelled to Japan and the rest of Asia, but it was not until the 15th century that it appeared in Europe (26) and North America had to wait until the mid 16th century (27). Today, tea is one of the most consumed beverages in the world (12, 28). While the role of tea in the development of cancer is still debated, its worldwide consumption is a sufficient reason to further define its association with lung cancer, one of the world leading causes of death.

2.2.1 Manufacturing

Tea comes from 2 varieties of a plant named Camellia sinesis; assamica and sinensis (27). The branches are regularly cut to harvest the new leaves used to produce all types of tea products (12). Tea can be generally separated into 3 types; black tea, green tea and oolong tea. Black tea is the most consumed worldwide and represents approximately 78% of the world consumption. Green tea is the most consumed type of tea in many parts of Asia and represents 20% of the world consumption, while oolong tea, principally consumed in specific regions of China, represents 2% of the world consumption (29). The three types of tea are manufactured in similar fashion with variation principally in the fermentation process and the resulting oxidative fermentation of the leaves (12). Black tea processing is divided into 5 steps: withering, rolling, fermentation, firing and grading (12). Withering includes the use of hot air to reduce moisture content in the leaves to 60%. During the rolling phase the tea leaves are crushed and macerated which leads to the oxidation of the flavonoids present in them. The fermentation is continued in the next step in which the leaves are spread into layers for up to 3 hours. The leaves are then passed through hot air driers to stop the fermentation and reduce the moisture content to 3%. Finally, black tea leaves are passed through screens and are graded to determine their quality before being shipped (12). The manufacture of green tea follows a similar procedure with the only variation being in the first step, where the leaves are heated using either the Japanese steaming method or the Chinese roasting method to prevent the oxidative fermentation process. After the heating, they are repeatedly rolled and dried until the moisture is reduced to 5%. The leaves are then roasted, ground and shipped, or shipped directly after the drying procedure to produce the various types of green tea (30). The manufacture of oolong tea is not as well standardized as green or black tea and tends to vary between factories. It is generally processed similarly to black tea but with only partial fermentation allowed (12).

2.2.2 Composition

Fresh tea leaves are composed of over a hundred chemicals (12) which can be categorised into a number of complex substances. The principal components are 25% of flavan-3-ols, 3% flavonols and flavonols glycosides, 5% phenolic acids and depsides, 3% of other polyphenols and 3% of caffeine per dry weight (27). The major flavan-3-ols (also known as catechins) present in fresh tea leaves are the following: epigallocatechin gallate (EGCG), epigallocatechin, epicatechin and epicatechin-3-gallate. Caffeine is present in similar amounts in the 3 main types of tea (27), however, the proportion of catechins, characterized by a "*di- or tri-hydroxyl group substitution of the B ring and the meta-5,7-dihydroxy substitution of the A ring*" (27), vary as a result of the oxidative fermentation required to produce black tea and oolong tea (27). Catechins represent 30% to 42% of the dry weight of green tea with EGCG representing 50% to 80% of all catechins (27, 31). By comparison, catechins only represent 3% to 10% of the dry weight of black tea. In black tea, flavan-3-ols and other phenolic compounds are oxidized to theaflavins, which represent 1% to 3% of the dry weight of black tea, and thearubigins which represent 10% to 40% of the dry weight of black tea and which are both absent from green tea (12). Various other types of polyphenol compounds account for 23% of the dry weight of black tea and 6% of the dry weight of green tea. Due to the various manufacturing methods of oolong tea, its composition is not as well defined, however the amount of the various substances is estimated to be between green and black tea (27, 31).

2.2.3 Mechanism for the hypothesized association with lung cancer

Several mechanisms by which tea intake could be associated with cancer risk have been proposed. The principal factors in both green and black tea that may be protective are catechins, which have strong antioxidant properties (31). EGCG is the most prevalent and bioavailable catechin in green tea (32). Animal and *in vitro* models have shown that catechins can increase cell cycle arrest, apoptosis, anti-inflammation, and anti-angiogenesis, and enhance detoxification and inhibit hypermethylation (32). While animal

models show strong support toward mechanisms for a protective effect of tea consumption on lung cancer, the results are less conclusive in humans. This could be due to a lower amount of tea present in the diet or because of a lower bioavailability of catechins in humans. While catechins are still present in black tea, a portion is oxidized into the less bioavailable but still protective theaflavin and thearubigen (32). The presence of caffeine in tea could result in a cancer promoting effect; however, this mechanism will be further discussed in section 2.3.3.

2.3 Coffee

Coffee is one of the oldest and most consumed beverages in the world. The plant first appeared in the Arabic peninsula from Ethiopia in the 6th century (12, 33) where the beans were eaten. It is not until the 15th century that coffee was consumed as a beverage named Kahwah in Yemen (12, 33). During the following century the Dutch imported coffee to their colonies and slowly became the principal producers and suppliers of coffee in Europe (33). The first European coffee shop opened at the end of the 17th century, around the same time coffee houses started appearing in the Americas (33). Today, over 130 million bags of coffee are consumed worldwide every year. Coffee represents 16% of the non-alcoholic beverage consumed in Canada, a proportion only equaled by soft-drinks (33, 34). In Quebec, 71% of adults consumed coffee on a daily basis in 2010 (35). Considering the worldwide rate of coffee consumption, it would be important to determine the role it plays in the development of lung cancer, one of the principal causes of death in the world.

2.3.1 Manufacture

Coffee comes from the Coffea genus of the Rubiaceae family of plants (12). Commercial coffee generally comes in two varieties; Arabica coffee which comes from C.arabica and Robusta coffee which come from C.canephora (12). Of the two types, Arabica is considered to be of higher quality and of better taste. As a result, Arabica coffee

represents over 75% of the world production (36). The fruit of the coffee plant is a cherry which take 3 to 4 years to grow before it can be harvested (37). Once it has been harvested, the cherry is quickly processed by either the dry method or the wet method. In the dry method the beans are dried until the moisture content is reduce to less than 11 percent. In the wet method, the beans are dried after being fermented in water-filled fermentation tanks. Once the beans are dry, they are milled, hulled, polished, graded and sorted by size (37). This "green" coffee is exported to be roasted at a temperature of 260 degrees Celsius which will give to the beans the known flavour of coffee. Finally, the coffee beans will be ground to different sizes depending on the type of coffee desired (37). Since the development of the first commercially viable instant coffee in Switzerland in 1938, coffee can be further processed to instant coffee, through a step that includes extraction, concentration, spray-drying, freeze drying and aromatization (12). Coffee beans can also be decaffeinated by either swelling the coffee beans with steam and water until their moisture content reaches 30%-40% in static or rotating drums with solvents, or by using supercritical carbon oxides in pressure vessels. The extracted caffeine is often used in the production of soft drinks and other caffeinated drinks (12).

2.3.2 Composition

A bean of roasted coffee is composed of a number of volatile and non-volatile compounds that vary in proportion based on the roasting method. On average, caffeine represents 1.3% of the dry weight of Arabica coffee and 2.4% of the dry weight of Robusta coffee (12). Polysaccharides such as lignins and pectins represent 33% and 37% of the dry weight of Arabica and Robusta coffee, respectively (12). Caramelized and condensation products such as Melanoidins represent 23% and 22.5% of the dry weight of Arabica and Robusta coffee respectively (12). Lipids represent 17% of the dry weight of Arabica coffee and 11% of the dry weight of Robusta coffee (12). Proteins represent 10% of the dry weight of both Arabica and Robusta coffee while the remaining dry weight is composed of minerals, triglonellines, acids, sugars and a number of non acid volatile substances (12). In

the case of instant coffee, up to 50% of the soluble substances could be extracted. The composition of a cup of coffee will vary greatly depending on the brewing method, brewing time and the amount of coffee used to prepare the beverage. In general, an estimated 75% to 100% of the caffeine content, 10% of the polysaccharides, 1% of the lipids, 20% to 25% of the caramelized or condensed products and 15% to 20% of the proteins are extracted from the beans (12). Considering a content of 10 g of coffee per 150 ml of water, a cup of filtered Arabica coffee would contain on average 102 to 120 mg of caffeine and a cup of filtered Robusta coffee would contain 187 to 220 mg of caffeine (12). A cup of espresso will generally contain less caffeine due to a lower extraction rate of caffeine (70% to 85%). As for instant coffee, the concentration of caffeine will vary depending on the type of coffee (Arabica or Robusta) and the amount used when making the beverage. It was estimated that in Canada the caffeine content of a cup of instant coffee (varying in size from 140 to 285 ml) varied from 29 to 176 mg/cup (12).

2.3.3 Mechanism for the hypothesized association with lung cancer

Caffeine has been shown to stimulate the proliferation of pulmonary adenocarcinoma and immortalized human small airway epithelial cells in *in vitro* experiments (13). It has been suggested that this is caused by the stimulation of the regulatory enzyme protein kinase A (PKA), the mitosis regulating extracellular signal-regulated protein kinase 1 and 2 (ERK1/2), and the cellular transcription factor cAMP response binding element (CREB) (13). In contrast, polyphenols present in coffee (e.g. chlorogenic acids derivatives) have been shown to reduce cancer proliferation through the inhibition of DNA methyltransferase, an enzyme which plays an important role in DNA methylation (38). Similarly, coffee as a beverage has been suggested to be an antioxidant in humans, which could reduce the risk of DNA damage (39).

3.0 Epidemiological studies

A review of the epidemiological evidence on the association between the consumption of tea and lung cancer, and on the association between the consumption of coffee and lung cancer was conducted using a combination of keywords in the Pubmed and Google scholar search engines. For tea, the keywords used were: tea, green tea, black tea, diet, caffeinated, beverage, lung cancer, cancer. For coffee, the keywords were: coffee, caffeinated, beverage, diet, lung cancer, cancer. There was no limit for the time of publication, however, only studies written in English or French that analyzed lung cancer in humans were included. Four studies (40-43), using the same study population, were not included in the review because they had been summarized and updated in 2008 (44). One written in Chinese (45) and two (46, 47) which analyzed lung, trachea and bronchus cancers together were also excluded from the review. In total, 25 studies published from 1988 to 2012 were included in this review (44, 48-71). A summary of the studies identified for the literature review on black tea (table 1) and coffee (table 2) can be found in appendix 1.

3.1 Tea

Fifteen studies analyzed the association between the consumption of black tea and lung cancer risk (44, 50, 52, 56-67). Of those, four were cohort studies (52, 57, 60, 65) and 11 were case-control studies (44, 50, 56, 58, 59, 61-64, 66, 67). Five of the studies did not indicate the type of tea consumed (57, 58, 61, 64, 66). However, they were conducted in countries in which black tea is the preferred type of tea (Finland (72), Canada (73), USA (74) and Sweden (75)), and we assumed that they analyzed black tea consumption. Because green tea is not the focus of this study, only a short overview of the literature will be given. Eleven studies analyzed the association between the consumption of green tea and lung cancer (44, 48-56, 68). One (56) described the effect of Okinawan tea, a slightly fermented

tea that is very similar to green tea. One (68) did not indicate the type of tea consumed, but since the study was conducted in China it was assumed that it was green tea. Of the 11 studies, eight were case-control studies (44, 49-51, 54-56, 68), and three where cohort studies (48, 52, 53).

3.1.1 Cohort studies

Cohort studies tended to report the presence of an inverse association between black tea consumption and lung cancer, although only half (two) of the studies (57, 65) reported statistically significant results. The first study, which analyzed a cohort of male smokers in Finland (57), reported a statistically significant inverse association when comparing subjects consuming at least 1 cup of black tea per day to subjects consuming less than 1 cup of black tea per day, with a RR (95%CI) of 0.66 (0.54 - 0.82). The second, analyzing a cohort of London males (65), calculated SMRs using the indirect standardization method and the population of England and Whale as the reference. They reported SMRs ranging from 0.63 to 1.41 for subjects consuming less than 4 cups, 4 to 6 cups, 7 to 9 cups and more than 10 cups of black tea per day. The results observed in this study were indicative of a u-shaped relationship between the consumption of black tea and lung cancer. This u-shaped association was not observed in a third cohort study (60), of Netherland residents or in a cohort of Hokkaido residents (52). However, in the Hokkaido study it was observed that the association varied by sex, with an RR of 0.6 (0.1 - 2.5) for males and an RR of 2.1 (0.3 - 17.5) for females, when comparing subjects consuming black tea weekly to subjects consuming black tea no more than monthly. It is important to note, however, that only 10 female cases were included in the analysis and large confidence intervals were observed for both males and females.

For green tea, only one (48) of the three Japanese cohort studies (48, 52, 53) reported a statistically significant inverse association between green tea consumption and lung cancer, with a RR of 0.57 (0.34 - 0.98) for female subjects consuming more than 9

cups of green tea per day, when compared to female subjects consuming less than 4 cups of green tea per day. No statistically significant association was reported for men. The other two cohort studies reported either statistically non-significant inverse associations (52) or statistically non-significant positive associations (53).

3.1.2 Case-control studies

The results reported in case-control studies, while inconsistent, tended to indicate the presence of an inverse association between the consumption of black tea and lung cancer. A statistically significant inverse association was reported by five of the 11 case-control studies (44, 56, 58, 61, 67), with ORs ranging from 0.10 to 0.69. Three of those studies (58, 61, 67) also reported an increase in the strength of the inverse association with an increase in the amount of black tea consumed daily. Three studies (59, 62, 66), reported principally statistically non-significant inverse associations while three studies (50, 63, 64) reported positive associations, among which one (63), analyzing Indian males, was statistically significant with an OR (95%CI) of 3.93 (2.55 - 5.92) when comparing ever black tea drinkers to never drinkers.

Two studies also conducted analyses by sex (44, 56). Statistically significant inverse associations were observed for female subjects consuming black tea compared to never drinkers, but not for male subjects. The validity of one of those studies (56) is of concern considering that all subjects were also consuming Okinawan green tea and only 5 female cases were included in the analysis. In the other study (44), the variation between males and females was only observed in non-smokers. Three studies only included males (63, 66, 67) while one study only included females (58). Statistically significant inverse associations between the consumption of black tea and lung cancer were reported in males (67), females (44, 56, 58) and both sex combined (61), while a statistically significant positive association was only reported in males (63).

Associations according to smoking status have also been reported in some studies (44, 61, 67). One study (44) reported a statistically significant inverse association for female non-smokers, but a statistically non-significant positive association for female ever smokers (p-interaction = 0.009). For male subjects, however, the interaction by smoking was not statistically significant. Another study (67) reported that while a statistically significant inverse association with black tea was observed for smokers, when subjects were analyzed based on the number of packs of cigarettes smoked per year the inverse association was stronger and only statistically significant in subjects smoking 44 or less packs of cigarettes per year. The third study (61) reported that while a statistically significant inverse association was observed when both smokers and non-smokers were analyzed together, when subjects were further analyzed based on their smoking status, the inverse association was only statistically significant for smokers. One study only included smokers (67), while two studies (58, 64) only included non-smokers. Statistically significant inverse associations were reported in non-smokers (44, 58) and smokers (67) and in both smoking status combined (56, 61). A statistically significant positive association was reported in both smoking status combined (63).

Only two studies (44, 59) examined the association according to histological type of lung cancer. One (44) examined adenocarcinoma cases, small cell carcinoma cases, and squamous cell carcinoma cases while the other (59) also examined large cell carcinoma cases. Both studies observed minimal variation in the ORs between the various histological types.

For green tea, the results from case-controls studies have been more consistent. A statistically significant inverse association was reported by six of the eight case-control studies (49, 51, 54-56, 68) with ORs ranging from 0.16 to 0.65. The two case-control studies (44, 50) which did not report any statistically significant associations between the consumption of green tea and lung cancer reported either weak statistically non-significant

inverse associations (50), or both weak statistically non-significant positive and statistically non-significant inverse associations (44).

3.3 Coffee

Thirteen studies analyzing the association between the consumption of coffee and the risk of lung cancer were included in this review (44, 52, 54, 58, 59, 62-64, 66, 67, 69-71). Of the 13 studies, three were cohort studies (52, 69, 71), 10 were case-control studies (44, 54, 58, 59, 62-64, 66, 67, 70) and two analyzed decaffeinated coffee (59, 62).

3.3.1 Cohort studies

Of the three cohort studies (52, 69, 71), two (69, 71) observed a statistically significant positive association between the consumption of coffee and lung cancer in males. One of the study (71) reported a statistically significant positive association between coffee consumption and lung cancer in a white American male cohort, with increasing RRs ranging from 2.1 to 2.4 for subjects consuming 3 to 4 cups, 5 to 6 cups and more than 6 cups of coffee per day when compared to subjects consuming less than 3 cups of coffee per day. When analysing the association by smoking status, it was reported that the positive association was mainly observed in current smokers. The second cohort study (69) reported, in a cohort of Norwegian residents, a statistically significant positive association with an HR of 2.4 for male subjects consuming 7 or more cups of coffee per day when compared to male subjects consuming 4 cups of coffee or less per day. For male subjects consuming 5 to 6 cups of coffee per day and for female subjects, the association was weaker and statistically non-significant. In a cohort of Hokkaido residents (52), a statistically non-significant positive association between the consumption of coffee and lung cancer was reported for female subjects with an RR of 2.1 (0.5 - 8.0). However, in male subjects a statistically non-significant inverse association with a RR of 0.7 (0.4 - 1.4) was suggested when comparing ever coffee drinkers to never coffee drinkers. Only 10

female cases were included in the later analysis and the difference observed between male and female subjects could be the result of chance.

3.3.2 Case-control studies

The associations between coffee consumption and lung cancer observed in case-control studies have been generally inconsistent. One cases-control study (70) reported a statistically significant inverse association, while one (63) reported a borderline statistically significant inverse association, with ORs for the two studies ranging from 0.41 to 0.69. Two studies (54, 59) reported statistically significant positive associations with ORs ranging from 1.51 to 1.82. Six (44, 58, 62, 64, 66, 67) reported only non-significant association of which three (44, 58, 64) reported only or principally inverse associations and three (62, 66, 67) reported only or principally positive associations. An increase in the strength of the positive association with an increase in the amount of coffee consumed was reported in one study (59), however, a stronger inverse association at higher amounts of coffee consumption has also been reported (58, 64). Moreover, two studies (66, 70) reported an inverse association at lower coffee consumption levels, but a positive association at higher coffee consumption levels. Two more (62, 67) reported no statistically significant differences in the strength of the association with an increase in the amount of coffee consumed.

Two studies (44, 54) conducted analyses by sex. One study (54) reported a statistically significant positive association for males but a statistically non-significant association for females. The second (44) observed similar results for males and females. Two studies only included females (58, 70), while three studies only included males (63, 66, 67). A statistically significant inverse association was reported in females (70), while a borderline statistically significant inverse association was reported in males (63). Statistically significant positive associations were reported in males (54) and in both sexes combined (59).

Associations according to smoking status have been reported only in one study (44), which reported no statistically significant interactions. Two studies only included non-smokers (58, 64), while 1 only included smokers (67). A statistically significant inverse association and a borderline statistically significant inverse association were reported in both smoking status combined (63, 70), as were the statistically significant positive associations (54, 59). However, one study (59) reported that when only non-smokers were included in the analysis, the observed ORs were slightly higher than in the main analysis.

Three studies (44, 54, 59) examined associations according to histological type of lung cancer. One study (44) reported similar associations between the consumption of coffee and lung cancer for cases with squamous cell carcinoma, small cell carcinoma or adenocarcinoma. The second study (57) reported that a moderate coffee consumption could be inversely associated with lung cancer in adenocarcinoma cases but not in large cell carcinoma cases, squamous cell carcinomas cases or small cell carcinoma cases. For the three later histological types, similar results were observed. The last study (54) reported stronger inverse associations and weaker positive associations for adenocarcinoma male cases when compared to small cell carcinoma male cases and squamous cell carcinoma male cases combined. For female cases, however, the statistical power was lower and similar results were observed among each histological type.

Two studies analyzed the association between the consumption of decaffeinated coffee and lung cancer (59, 62), and both reported statistically significant inverse associations, with ORs ranging from 0.44 to 0.67 when comparing subjects consuming decaffeinated coffee to subjects not consumed the beverage.

3.4 Summary of evidence

While the evidence supporting an association between the consumption of black tea and lung cancer risk is weaker than for green tea, there have been indications that black tea

may be inversely associated with lung cancer. One cohort study (57) and five case-control studies (44, 56, 58, 61, 67) observed statistically significant inverse associations between black tea consumption and lung cancer, while only one case-control study (63) observed a statistically significant positive association, and one cohort study (65) observed statistically significant inverse associations at consumption of 6 cups of black tea per day or less, but statistically significant positive associations for higher consumption levels. There were some indications that the association between the consumption of black tea and lung cancer varied by sex (44, 52, 56) and smoking (44, 61, 67). However, the evidence was weak due to low statistical powers and contradictory observations. The association between the consumption of black tea and lung cancer has not varied by the histological type of the cancer.

For coffee, cohort studies tended to report statistically significant positive associations between the consumption of coffee and lung cancer (69, 71). However, results from case-control studies have been inconsistent, with statistically significant positive associations (54, 59), a statistically significant inverse association (70), and a borderline statistically significant inverse association (63) reported. There has been a tendency for positive associations to be only observed at higher levels of coffee consumption (54, 59, 66, 69-71), and in some cases a significant linear trend was observed (59, 69, 70). There were weak indications that the association between coffee consumption and lung cancer varied by sex (52, 54, 69), smoking status (71) and that it was more strongly inversely associated with adenocarcinoma (54, 59). Two studies analyzed decaffeinated coffee (59, 62) and both observed strong statistically significant inverse associations. Because for both studies the reference group included caffeinated coffee drinkers, those results suggest that caffeine is a risk factor for lung cancer. However it is also possible that the observed association is confounded by healthier life habits amongst subjects consuming decaffeinated coffee.

Limitations in the assessment of the exposure could be responsible for the inconsistencies of the literature; lung cancer take decades to develop, however, the assessment of the consumption of black tea and coffee was generally limited to a few years preceding the interview and therefore may not have captured the history of consumption of those beverages during the etiologically relevant period of time. Moreover, patients can change diet habits, including beverages consumption, in the years preceding the diagnosis of lung cancer because of clinical manifestations preceding the disease, which can result in a reverse causality bias in case-control studies. The duration of consumption might also play an important role in the association between those two beverages and lung cancer, but to our knowledge this aspect of the consumption has never been analyzed. Seven of the 15 studies on black tea included less than 500 cases, while only six of the 13 studies on coffee included more than 300 cases, which reduced the statistical power of the analyses. There were also variations in the definition of a cup of black tea or coffee and in the categorization of the main variable that could be partially responsible for the inconsistencies of the literature. For example, a majority of studies which reported a statistically significant positive association between coffee consumption and lung cancer used light to never drinkers of coffee as the reference group, while studies reporting a significant (or borderline significant) inverse association tended to only use never drinkers.

4.0 Research objectives

The objectives of this research are to examine the association between different aspects of the consumption of black tea and of the consumption of coffee, and lung cancer risk.

5.0 Methodology

5.1 The study

Analyses were done using data obtained from the Montreal lung cancer case-control study (MLCS). The MLCS study period was from 1996 to 2001. The study population included Canadian citizens aged 35 to 75 years old and resident of the island of Montreal, the South-Shore of Montreal, and Laval, representing 3.1 million people. The study was approved by the research ethics boards of McGill University, the Armand-Frappier institute and the 18 hospitals who agreed to participate in the study.

5.1.1 Cases

Cases were individuals diagnosed between 1996 and 1997 with a primary lung cancer and were selected from one of 18 major hospitals serving the study region, representing 98% of all declared lung cancer cases in the region. The study staff was informed of all new cases of lung cancer by staff members from hospital tumour registries, pathology departments or the medical records departments from the hospitals. When a case satisfied the age and residence criteria, the case physician would be informed and asked for authorization to contact the subject. 45 cases could not be contacted because the physician refused to give the authorization. Of 1429 contacted cases, 1203 accepted to participate for a response rate of 84.2%.The response rates were slightly higher among males (83.4%) than females (81.3%). A total of 738 male and 465 female cases accepted to participate in the study. Each case was histologically defined according to the classification of the International Agency for Research on Cancer.

5.1.2 Controls

Controls were selected from the Quebec voter registration list and were frequency matched to the distribution of the cases by 5 year age group, sex and electoral district. To ensure that each case would have a control, 3 "attempted" controls were selected for each case. If the first control was not eligible the next in line would be selected and the procedure would be repeated until a control accepted to participate. Of the 2179 controls contacted, a total of 1513 accepted to participate, for a participation rate of 69.4%. The participation rate was similar among males (69.5%) and females (69.2%), for a total of 899 males and 614 females who accepted to participate in the study.

5.2 Exposure assessment

5.2.1 Data collection

Each participant was first sent a letter requesting participation, along with a short self-administered questionnaire. The questionnaire was used to confirm the address, phone number, the city or country and date of birth, and if applicable, the date of entry into Canada. A brief section on occupational history was also included and used by the interviewers to prepare for the in-depth occupational section of the interview. Face-to-face interviews were conducted by interviewers trained specifically for this study and lasted an average of 2 hours. Interviewers were not blinded to the status of the subjects. The informed consent of each participant was obtained before the interview. The first part of the interview was a structured questionnaire which gathered information on non-occupational factors, including socio-demographic factors, active and passive smoking, intake of tea, coffee, and alcoholic beverages, well water consumption, as well as cooking and heating methods. A semi-structured work history questionnaire was then used for the rest of the interview. The goal of this section was to obtain detailed information on any work held for more than a year by the participant. The questionnaire included questions on the responsibilities held by the participant, ergonomic factors, and the presence of known

carcinogens in the work environment. To better assess the exposure history of specific occupations such as welding and farming, 31 specialized questionnaires were used to complement the main exposure questionnaire. During the first few months of the study a second self-administered questionnaire was given to each participant at the end of the interview, to be sent back to the study team after completion. However, given the low return rate, it was decided to integrate this second self-administered questionnaire into the face-to-face interview. This questionnaire asked further questions on factors related to the home environment, environmental tobacco smoke, and physical activities. Additionally, dietary habits for the past two years were assessed using a semi quantitative food frequency questionnaire (FFQ) developed by the Canadian Cancer Registries Epidemiologic Research Group. The FFQ was composed of 42 items, including fruits, vegetables, grain products, dairy products and meats. A proxy replaced the participant during the interview for 37.7% of cases and 7.8% of controls.

The structured interview included 6 questions (figure 1) on the consumption of tea and coffee. Question 48A and 48B assessed whether subjects ever drank tea or coffee at least once a week, and if they did, whether they drank it nearly every day. For those participants that drank tea or coffee nearly every day, questions 48C to 48F assessed their duration of consumption by asking for the age at which they had started their consumption, whether they were still consuming the beverage and if not, the age at which they had stopped their consumption. The last question, question 48F, asked subjects to indicate their average daily consumption (in cups) of tea or coffee.

Subjects were not asked for the type of tea or coffee consumed. Regarding tea, it is known that black tea has traditionally been the preferred type consumed in Canada since its introduction in the country (73). Moreover, import of green tea in bags has grown by 50% only since the late 90's to represent 17% of total tea import in 2009, indicating that less than 10% of tea imported and consumed during and before the 90's was green tea (76).

Taking these consumption patterns into consideration, it was assumed that participants in our study consumed black tea. With regard to decaffeinated tea, while its consumption has become more popular since the 90's, the worldwide rate of decaffeinated tea consumption is generally considered to be low (12), with a reported rate of less than 10% in the United-states in 2000 (77). It was assumed that participants in our study consumed principally caffeinated black tea.

For coffee, several factors contribute to the type consumed, including personal preferences and place of consumption (home, coffee shop, restaurant). Moreover, the consumption rate of one of the principal source of Robusta coffee, instant coffee, varied from as much as 52% of all coffee consumed in Canada during the late 80's (12) to as low as 13% in 2010 (35). Our study did not differentiate between the various types of coffee. Given the possible role of caffeine in the development of lung cancer, it would be important to know the rate of consumption of decaffeinated coffee in our study population. The proportion of decaffeinated coffee consumed worldwide has shifted in the past, with a peak in popularity, particularly strong in the United States, during the late 80's (78). Still, even during those years the rate of decaffeinated coffee consumption was never particularly high with reported rates of 6% in the United Kingdom, 2.2% in Italy, 13.8% in Germany and 16.7% (12) in the United-States. Moreover, decaffeinated coffee started to lose popularity during the 90's and was estimated to represent around 10% of worldwide sales in 2010 (78). Given that our questionnaire did not include questions on the type of coffee consumed, it was assumed that participants in our study population consumed caffeinated coffee.

5.3 Statistical analysis

5.3.1 Subjects included in the analysis

To be included in the analysis, subjects had to have complete information on either their black tea or coffee consumption and on their smoking history. The definition of complete information was for subjects to have indicated if they had ever consumed black tea or coffee (Q48A in Appendix II). Subjects who had ever consumed one of the beverages had to have indicated if they drank it nearly every day (Q48B) and if they did, they had to have answered all remaining questions (Q48C to Q48F). 72 cases and 23 controls were excluded from the analysis due to missing information on both their coffee and black tea consumption. 8 more subjects, 1 case and 7 controls, were excluded due to missing information on their smoking habits. Of these 103 subjects with missing information, 64 cases and 15 controls did not directly participate in the interview but instead had a proxy respond for them. In total, 1130 cases (93.9% of interviewed cases) and 1484 controls (98% of interviewed controls), were included in the analyses of tea and/or coffee intake.

The analyses of black tea included 1111 cases (98.3% of included cases) and 1469 controls (99.1% of included controls) that had complete information on their black tea consumption. Of the 34 subjects without complete information, 17 cases and 7 controls were represented by a proxy during the interview. For coffee, 1009 cases (89.3% of included cases) and 1434 controls (96.7% of included controls) had complete information on their coffee consumption and were included in the analysis. Of the 171 subjects without complete information on their coffee consumption, 99 cases and 30 controls were represented by a proxy during the interview. As could be expected, the proportion of subjects represented by a proxy during the interview was high for subjects with missing information on their black tea or coffee consumption. More subjects had complete information for their black tea consumption than for their coffee consumption. This could be related to proxies. The consumption of black tea, and particularly of coffee, can occur both at home and outside of home, at any time of the day. It can be complicated for a proxy

to know the exact consumption rate of an ever drinker, however, this is less of an issue if the subject never consumed the beverage. Since more subjects were never drinkers of black tea (45%) than coffee (7.2%), it is probable that proxies had more difficulty on reporting coffee consumption of the subject they represented than on reporting their black tea consumption. A flowchart of the selection process can be found in figure 2 (appendix 3).

5.3.2 Dependent variable

The dependent variable is a binary variable based on the disease status of the subjects and was coded as either 1 for cases of lung cancer or 0 for controls.

5.3.3 Main independent variables

5.3.3.1 Frequency and average daily amount of black tea and coffee consumption

Two categorical variables were created to examine the association between the intensity of consumption of black tea and lung cancer: one variable for the frequency of consumption and the other for the average daily amount of consumption. The variable for the frequency of black tea consumption included three categories; “never drinkers”, “occasional drinkers” and “regular drinkers”. Occasional drinkers were defined as subjects consuming black tea at least once a week but not nearly every day, while regular drinkers were defined as subjects consuming black tea nearly every day. The variable for the average daily amount of consumption of black tea was categorized as “never drinkers”, “less than 1 cup per day”, “1 cup per day”, “more than 1 cup to 3 cups per day” and “more than 3 cups per day”. Subjects defined as occasional drinkers were not asked for their average daily amount of consumption. In this analysis, it was assumed that their average daily amount of consumption was less than 1 cup per day and they were categorized as such. All subjects, with the exception of 7 controls, that were categorized as consuming less than 1 cup of black tea per day were occasional drinkers.

For coffee, the frequency of consumption variable was created similarly to black tea and was composed of the same 3 categories. For the average daily amount of consumption variable, however, higher consumption of coffee allowed for more categories. The categories were: "never drinkers", "less than 1 cup per day", "1 cup per day", "more than 1 cup to 3 cups per day", "more than 3 cups to 5 cups per day" and "more than 5 cups per day". As for black tea, occasional coffee drinkers were assumed to have an average daily amount of coffee consumption of less than 1 cup of coffee per day. All subjects, with the exception of 1 case and 6 controls that were categorized as consuming less than 1 cup per day were occasional drinkers.

5.3.3.2 Duration of consumption of black tea and coffee

The association between the duration of black tea consumption and lung cancer was analyzed using a categorical variable. The duration of consumption was first calculated by subtracting the age at which subjects indicated they had started consuming the beverage from either the age at which they had stopped their consumption or if they were still consuming the beverage, the age at the time of the interview. The duration of consumption was categorized into six categories. Because only a small amount of black tea or coffee is present in each cup ingested, the continuous consumption of the beverage over a long time period could be required for a change in the association with lung cancer to be observed. As such, it was decided to categorize subjects in 10 year increments. Because too few subjects consumed black tea for less than 10 years, subjects who had consumed black tea for less than 20 years were categorized together. The same was true for subjects who had consumed black tea for 50 years or more. The final categories for the duration variable were: "never drinkers", "less than 20 years", "20 years to less than 30 years", "30 years to less than 40 years", "40 years to less than 50 years" and "50 years or more". The duration variable was coded similarly for coffee. Health issues associated with advanced stages of lung cancer, such as fatigue, chest pain, nausea and difficulties to swallow (17) can change life habits of cases preceding the diagnosis of the illness. As such the drinking status of subjects during

the two years preceding the interview were judged as non-representative of the consumption of those beverages during the etiologically relevant period of time and were removed from the analysis for both cases and controls. Because occasional drinkers were not asked questions regarding the duration of their consumption, they were not included in this analysis.

5.3.3.3 Cumulative consumption

To analyse the cumulative consumption of black tea, both the duration and the average daily amount of consumption of black tea were used to create a new variable. Special care was taken to create categories that would be representative of the variation in the average daily amount of consumption and in the duration of consumption observed in our subjects, while insuring that a sufficient number of participants would be present in each category to obtain precise results. Subjects were first categorized into one of three categories based on their average daily consumption: “never drinkers”, “1 cup or less per day” and “more than 1 cup per day”. Among the latter two categories, subjects were further categorized into 3 categories based on their duration of consumption: “less than 30 years”, “30 years to less than 50 years” and “50 years or more”, resulting in a total of 7 categories. This categorization allowed for the analysis of a light and heavy consumption of black tea over a long, medium and short duration of consumption. A similar variable was created for coffee.

5.3.4 Potential covariates

Potential covariates were selected based on the possibility that they could be related to both the consumption of black tea or coffee and lung cancer risk. Information on potential covariates came principally from the literature review on the association between the consumption of black tea or coffee and lung cancer. Any potential covariate for which data was available was considered. Age, sex and ethnicity were selected due to both their

use in the literature and from the past experience of working with the MLCS data. Originally, subjects were classified in one of 14 ethnic groups. However, 84% of cases and 72% of controls were either French or English Canadian. As a consequence, the remaining 12 categories were merged resulting in a total of 3 categories: French Canadian, English Canadian and other ethnicities.

5.3.4.1 Socio-economic factors

Socio-economic factors have been associated with lung cancer (3), and diet has been shown to vary between socio-economic groups (79-81). As such, two potential socio-economic covariates were considered for the analysis: income (median census tract family income), and education. A census tract was defined as a small geographical area with a population of 2500 to 8000 individuals (82). Income was categorized into 3 categories: low income, middle income and high income, based on the tertiles of the controls. Education was represented by number of years of schooling and was categorized into 3 categories; less than 7 years, 7 to less than 12 years and 12 years or more, approximately representing up to primary school, up to secondary school and higher education respectively.

5.3.4.2 Respondent status

In some situations subjects could not present themselves to the interview, in which case a proxy replaced them. Proxies are not as reliable as subjects and are more prone to error during the interview. Since more cases than controls were represented by proxies in our study, this could result in a differential misclassification error. As such, the respondent status was included as a potential covariate to change this differential misclassification error to a non-differential one.

5.3.4.3 Comprehensive smoking index

Smoking is one of the most important risk factors of lung cancer (2) and has been associated with both diet in general (80) and tea or coffee consumption in particular (83-85). It is very important to adjust for smoking as thoroughly as possible in this study. To do so, the comprehensive smoking index (CSI) developed by Leffrondre et al (86) was used. This index originally proposed by Hoffman et al (87), was modified for lung cancer in our study and fitted using the data from the MLCS, and is a measurement of the cumulative exposure to smoking. Its formula is composed of three terms that include, in one measure, the duration of smoking (years) of the subject, the time since the cessation of smoking (years) and the natural logarithm of the average intensity of smoking (in cigarette per day). This index has been shown to be one of the most complete and effective ways to model and control for the confounding of smoking in the MLCS (86).

5.3.4.4 Dietary factors

Intake of fruit and vegetable, and to a lesser extent alcohol have all been associated with lung cancer (3, 24, 88, 89), and with tea and coffee consumption (84, 90). As such, fruit intake, vegetable intake, as well as intake of beer, wine and spirits were considered as potential covariates. Beer, wine and spirits were categorical variables with three categories defined similarly to the frequency of consumption variable of black tea; “never drinkers”, “occasional drinkers” and “regular drinkers”. Vegetable and fruit intake were categorical variables with 4 categories based on our controls quartiles of intake. To reduce the measurement error of the diet factors due to confounding by total energy intake (91), total energy intake was included in the model as a continuous variable in Kcals/day. It is important to note however, that the FFQ did not elicit information on all possible foods, thus energy intake was underestimated in our study population. Black tea was also considered as a potential covariate for coffee and coffee for black tea. In both cases the frequency of consumption variable was used.

5.3.4.5 Occupational exposures

Exposure to occupational agents such as asbestos has been associated with lung cancer (3), and while no direct association has been shown to exist between exposure to specific occupational agents and the consumption of black tea or coffee, an association has been shown to exist between unhealthy diet habits and blue collar occupations (79, 81). Since jobs associated with exposure to occupational agents tend to be blue collars occupations, exposure to occupational agents could be a potential covariate in this study. Detailed information on the lifetime exposure to nearly 300 occupational agents was available for each subject following their assessment by a team of industrial hygienists and chemists (92). The team used three measurements; the concentration of the agent present at work (low, medium, high), the frequency of exposure in a normal work week (percentage of the week time), and the reliability that the exposure had occurred (possible, probable and definite) to determine the lifetime exposure to each agent. The lifetime exposure was coded as a categorical variable with three categories: no exposure, nonsubstantial exposure, and substantial exposure. For each occupational agent two measurement variables were available, one including all reported exposures and one only including exposures considered as reliable. Using the latter variable, exposure to asbestos and exposure to silica, two occupational agents both established to be associated with lung cancer (3, 21, 22) and with sufficient amount of exposed subjects in our study population, were recoded as binary variables (exposure, no exposure), and included as potential covariates in the model.

5.3.4.6 Model reduction procedure

Age, sex, income, smoking, respondent proxy status and education are potential covariates well described in the literature on the association between black tea or coffee and lung cancer. Thus, they were "forced" into the model. While ethnicity was not as well described, it was associated with lung cancer and with the consumption of black tea and coffee and was also "forced" in the model. For vegetable, fruit, beer, wine and spirit intake as well as for exposure to asbestos and exposure to silica, the evidence was weaker and a

model reduction procedure was used to determine their inclusion in the model. Since the inclusion of the energy intake variable was dependent on the inclusion of at least one diet variable, it was not tested. For this procedure, the variable for the frequency of consumption of black tea or coffee was used. The base model was adjusted for age, sex, ethnicity, median census tract family income, total years of education, respondent proxy status, smoking (CSI), vegetable intake, fruit intake, beer intake, spirit intake, wine intake, energy intake, exposure to silica and exposure to asbestos. Then, one at a time, each potential covariate was removed from the model and the percentage change in the OR for the frequency of consumption variable was calculated. The covariate whose exclusion resulted in the smallest change in the OR was removed and then the model was run once again. The process was repeated until all suspected covariates were examined. If removal of a potential covariate from the model resulted in a change in the estimated OR of 10% or more then it was kept in the final model. A summary of the model reduction procedure for black tea and coffee can be found in tables 3 and 4 (appendix 4). For black tea, no suspected covariates modified the OR by at least 10%, however wine intake modified the OR by 9.4% and it was decided that the modification was strong enough to keep the variable in the final model. For coffee, only wine intake modified the OR by at least 10% and was kept in the final model. To test black tea as a potential covariate for coffee, and coffee as a potential covariate for black tea, a new study sample including only subjects with complete information for both beverages was created. Using this study population the model reduction procedure was used to determine the inclusion or exclusion in the model of those two potential covariates (results not shown). However, since they modified the OR by less than 5% they were not kept in the final model. The final model was adjusted for age, sex, ethnicity, median census tract family income, total years of education, respondent proxy status, smoking (CSI), wine intake and energy intake.

5.3.5 Missing values

Missing values were present for 2 covariates: wine intake and energy intake. To avoid further reducing the amount of subjects present in the analysis, values were given to replace the missing information using the single imputation method. While contested (93) this method has been shown to be effective when the number of subjects with missing values is small (94, 95). Using this method the 9 cases and 2 controls that did not indicate their wine intake and the 51 cases and 41 controls that did not indicate their energy intake were assigned the value of the median among controls.

5.3.6 The logistic regression model

The logistic regression model, derived from the logistic function $f(z) = 1/(1+e^{-z})$, where e is the natural logarithm, z, which can vary from $-\infty$ to ∞, represent a linear combination of independent variables and f(z), the expected value of the outcome, can vary from 0 to 1, is the most commonly used model for the analysis of the association between multiple independent variables and a binary dependent variable (96). The model allows for the estimation of the odd ratios (OR) that the outcome would occur at different levels of exposure to a main variable given that the exposure to the rest of the variables is constant. For this study, it allowed the comparison of the risk of developing lung cancer between subjects with various frequencies of consumption, average daily amounts of consumption, durations of consumption and cumulatives consumption of black tea or coffee, and subjects who never consumed the beverage, while adjusting for potential covariates. As an indication of the precision and statistical significance of the estimates, the 95% confidence intervals (95%CI) of the ORs were estimated.

As an indication of the presence of a linear dose-response between the average daily amount of consumption and the duration of consumption of black tea or coffee and the risk of lung cancer, a test for trend was carried out. The p-value for the test was obtained by entering the variable as a continuous term in the logistic regression model and assessing the

Wald chi-square test statistic. We also conducted a test (named "global test" for this study) of the overall significance of the association observed in each model. This test was made using the likelihood ratio (LR) test. The LR test is calculated using the formula LR = (-2 log L1) – (-2 log L2) where L1 is the maximum likelihood of model 1 (the model without the main variable) and L2 is the maximum likelihood of model 2 (model including the main variable). The LR test follows a chi-square distribution in large samples with the difference in the number of parameters between the two models used as the degrees of freedom.

5.3.7 Main analysis

Each of the four main exposure variables for black tea and coffee intake, i.e. (1) the frequency of consumption, (2) the average daily amount of consumption, (3) the duration of consumption, and (4) the cumulative consumption were analyzed in separated multivariate models using unconditional logistic regression in the Statistical Analysis Software 9.0.0 (SAS). For each of the 8 analyses, the reference group was never drinkers.

5.3.8 Secondary analyses

5.3.8.1 Stratification by sex

There were some indications in the literature that the OR for the association between the consumption of black tea or coffee and lung cancer varied by sex (52, 54, 56). In order to analyse the role of gender as a potential effect-measure modifier for the association between the consumption of black tea or coffee and lung cancer, analyses stratified by sex were conducted. Because stratification reduces the number of subjects present in each analysis, which in turn reduces the precision of the results, the duration variable had to be reduced to 5 categories: "never drinkers", "less than 30 years", "30 years to less than 40 years", "40 years to less than 50 years" and "50 years or more". We computed the power for detecting the interaction by sex for black tea drinking (never

drinkers/ever drinkers) based on the following inputs: 1111 cases, 1469 controls, 40% of male subjects, 45% of ever drinkers of black tea, an α of 0.05, an OR (I.E. the ratio of the OR for black tea amongst females compared to the OR of black tea amongst males) for the interaction by sex of 1.5, an OR of 0.8 for ever drinkers of black tea (vs never drinkers) and an OR of 1 for female subjects (vs male subjects). A statistical power of 70% was obtained. For coffee, the power was calculated based on the following inputs: 1009 cases, 1434 controls, 40% of male subjects, 93% of ever drinkers of coffee , an α of 0.05, an OR for the interaction by sex of 1.5, an OR of 1.4 for ever drinkers of coffee (vs never drinkers) and an OR of 1 for female subjects (vs male subjects) . A statistical power of 21% was obtained.

5.3.8.2 Stratification by smoking

There were some indications in the literature that the association between the consumption of black tea or coffee and lung cancer varied between smokers and non-smokers (44, 61, 67, 71). To further define the role of smoking in the association between the consumption of black tea or coffee and lung cancer, analyses stratified by smoking were conducted. In order to run the stratified analyses, subjects were categorized based on their CSI. Originally, the analyses were intended to have three strata: “non-smokers”, “light smokers” and “heavy smokers”. Light smokers were defined as smoking subjects with a CSI equal as or less than the median of 1.928, while heavy smokers were subjects with a CSI higher than the median. Because only 47 cases were non-smokers, it was decided to merge light smokers and non-smokers into one stratum. Each stratum was still adjusted for smoking using the CSI variable. As was the case for the stratification by sex, the duration variable was reduced to 5 categories to increase the precision of the analysis. We computed the power for the interaction by smoking for black tea drinking using the methodology described in the previous section based on the following inputs: 1111 cases, 1469 controls, 40% of heavy smoker subjects, 45% of ever drinkers of black tea, an α of 0.05, an OR for the interaction by smoking of 1.5, an OR of 0.8 for ever drinkers of black tea (vs never

drinkers) and an OR of 10 for Heavy smokers (vs light smokers). A statistical power of 50% was obtained. For coffee, the power was calculated based on the following inputs: 1009 cases, 1434 controls, 40% of heavy smoker subjects, 93% of ever drinkers of coffee, an α of 0.05, an OR for the interaction by smoking of 1.5, an OR of 1.4 for ever drinkers of coffee (vs never drinkers) and an OR of 10 for Heavy smokers (vs light smokers). A statistical power of 17% was obtained.

5.3.8.3 Analysis by histological type

The association between the consumption of black tea and coffee was further explored in analyses that considered the histological type of the cases. In the MLCS, lung cancer cases were categorized in one of six histological types; squamous cell carcinoma, small cell carcinoma, adenocarcinoma, large cell carcinoma, other epithelial tumors and carcinoma not otherwise specified. Of those, only three types had enough cases to be analyzed: Squamous cell carcinoma, small cell carcinoma and adenocarcinoma. Each histological type was analyzed separately and included the cases with the specified type and all controls. As was the case for the stratified analyses, the duration variable was reduced to 5 categories.

6.0 Results

6.1 Selected characteristics of the study population

The distributions of the selected characteristics of the study population are shown in table 5. A majority of subjects were male (60%) and more than half (51%) were over 65 years old. There were more French Canadian cases (78%) than controls (66%), while more controls (29%) than cases (16%) came from ethnicities other than French or English Canadian. Controls were more educated and had a higher median census tract family income than cases, with 41% of controls compared to 26% of cases having studied for 12 years or more, and 33% of controls compared to 23% of cases having a high median census tract family income. Smoking was more prevalent for cases with 96% of the cases having smoked (CSI > 0) compared to 69% of controls. Cases had a lower wine intake than controls with 15% and 20% of cases consuming wine daily or weekly, respectively, compared to 18% and 37% among controls. Similarly, cases also had a lower energy intake than controls with close to half of the cases (49%) consuming less than 800 Kcals per day compared to a third (33%) of controls. The use of a proxy during the interview occurred for 34% of the cases and 7% of the controls. Adenocarcinoma (39%) and squamous cell carcinoma (29%) were the most prevalent histological types of lung cancer among cases.

Table 5: Selected characteristics of the study population

	Cases (n=1130)	Controls (n =1483)
Age		
<45 years	37 (3.3)	52 (3.5)
45 - <55 years	164 (14.5)	202 (13.6)
55 - <65 years	381 (33.7)	442 (29.8)
65 - <75 years	511 (45.2)	745 (50.3)
≥75 years	37 (3.3)	42 (2.8)
Sex		
Male	677 (59.9)	878 (59.2)
Female	453 (40.1)	605 (40.8)
Respondent status		
Self	741 (65.6)	1380 (93.1)
Proxy	389 (34.4)	103 (6.9)
Ethnic group		
French Canadian	876 (77.5)	979 (66.0)
English Canadian	74 (6.6)	82 (5.5)
Other	180 (15.9)	422 (28.5)
Years of schooling		
<7	285 (25.1)	312 (21.0)
7 - <12	553 (48.9)	563 (38.0)
≥12	292 (25.8)	608 (41.0)
Mean census tract family income		
Low	500 (44.3)	491 (33.1)
Middle	369 (32.7)	503 (33.9)
High	261 (23.1)	489 (33.0)
Cigarette smoking index		
0	47 (4.2)	460 (31.0)
>0 - ≤2	388 (34.3)	738 (49.8)
>2	695 (61.5)	285 (19.2)
Wine consumption		
Never	739 (65.4)	665 (44.8)
weekly	25 (19.9)	545 (36.8)
Daily	166 (14.7)	273 (18.4)
Daily energy intake (Kcal)[a]		
<800	551 (48.8)	495 (33.4)
800 - <1000	301 (26.6)	475 (32.0)
1000 - <1200	167 (14.8)	329 (22.2)
≥1200	111 (9.8)	184 (12.4)
histological type		
Squamous cell carcinoma	328 (29.0)	
Adenocarcinoma	443 (39.2)	
Small cell carcinoma	187 (16.6)	

a: Daily energy intake calculated from the 42 items of the food frequency questionnaire

6.2 Black tea

6.2.1 Analysis of the frequency and average daily amount of consumption

The unadjusted ORs (95%CI) indicated that occasional and regular black tea consumption (table 6) was statistically significantly associated with a reduced risk of lung cancer, with a stronger protective association with occasional consumption (OR: 0.50 (0.39-0.63)) than with regular consumption. However, after adjustment for several covariates, associations were no longer significant. While an inverse association was still observed for an occasional consumption of black tea (OR: 0.80 (0.59 - 1.08)), a positive, albeit statistically non-significant, association was observed for regular consumption (OR: 1.08 (0.87 - 1.34)).

Table 6: OR estimates for the frequency of black tea consumption in relation to lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95%CI	OR	95% CI
Never	551	604	1.00	ref	1.00	ref
Occasional	127	281	0.50	0.39 – 0.63	0.80	0.59 - 1.08
Regular	433	584	0.81	0.69 – 0.96	1.08	0.87 - 1.34
p-value (global)				< 0.01		0.16

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

A similar pattern was observed when examining the average daily amount of consumption (table 7), though the unadjusted estimates suggested an increased risk with consumption of ≥ 3 cups per day (OR: 1.39 (0.97 – 1.99)). However, no statistically significant association was observed in the multivariate model. The statistically non-significant positive association observed for a regular frequency of consumption of black tea was only observed at consumption of > 1 to 3 cups/day (OR: 1.15 (0.88 – 1.50)) and > 3 cups /day (OR: 1.11 (0.72 – 1.73)).

Table 7: OR estimates for the average daily amount of black tea consumption in relation to lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI
Never	551	604	1.00	Ref	1.00	Ref
< 1 cup/day	127	288	0.48	0.38 - 0.61	0.78	0.58 - 1.05
1 cup/day	128	231	0.61	0.48 - 0.78	1.00	0.73 - 1.35
>1 to 3 cups/day	229	286	0.88	0.71 - 1.08	1.15	0.88 - 1.50
>3 cups/day	76	60	1.39	0.97 - 1.99	1.11	0.72 - 1.73
p-value (trend)				0.69		0.29
p-value (global)				< 0.01		0.26

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

6.2.2 Analysis of the duration of consumption

While only inverse associations were observed in the unadjusted model (table 8), the pattern of adjusted point estimates implied a possible non-monotonic association. Statistically non-significant positive associations were observed for black tea consumption of < 20 years (OR: 1.46 (0.90 – 2.36)) and of ≥ 50 years (OR: 1.28 (0.90 – 1.83)) relative to never drinkers. For consumption of 20 years to < 50 years, statistically non-significant ORs ranging from 0.94 to 1.01 were observed.

Table 8: OR estimates for the duration of black tea consumption in relation to lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI
Never	551	604	1.00	Ref	1.00	Ref
< 20 years	52	61	0.93	0.63 - 1.38	1.46	0.90 - 2.36
≥ 20 to < 30 years	39	55	0.78	0.51 - 1.19	1.01	0.60 - 1.72
≥ 30 to < 40 years	80	106	0.83	0.61 - 1.13	0.99	0.67 - 1.46
≥ 40 to < 50 years	141	203	0.76	0.60 - 0.97	0.94	0.69 - 1.29
≥ 50 years	121	159	0.83	0.64 - 1.09	1.28	0.90 - 1.83
p-value (trend)				0.02		0.60
p-value (global)				0.24		0.46

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

6.2.3 Analysis of the cumulative consumption

There was a tendency for the unadjusted estimates to be closer to the null value for subjects consuming > 1 cup/day when compared to subjects consuming ≤ 1 cup/day (table 9). For the adjusted estimates, non-monotonic associations similar to what was observed in the adjusted duration of consumption analysis were observed for both a consumption of ≤ 1 cup/day and > 1 cup/day. However, with the exception of subjects who had consumed black tea for 30 years to < 50 years, for whom more variation was observed, the ORs only varied by 4% or less between subjects with different average daily amounts of consumption.

Table 9: OR estimates for the cumulative consumption of black tea in relation to lung Cancer

				Crude		Adjusted*	
		Cases	Controls	OR	95% CI	OR	95% CI
Never	-	551	604	1.00	Ref	1.00	Ref
≤ 1 cup/day	< 30 years	36	56	0.71	0.46 - 1.09	1.26	0.73 - 2.19
	≥ 30 to < 50 years	56	131	0.47	0.34 - 0.65	0.76	0.51 - 1.14
	≥ 50 years	36	51	0.77	0.50 - 1.20	1.24	0.70 - 2.21
> 1 cup/day	< 30 years	55	60	1.01	0.69 - 1.48	1.22	0.77 - 1.94
	≥ 30 to < 50 years	165	178	1.02	0.80 - 1.29	1.08	0.79 - 1.46
	≥ 50 years	85	108	0.86	0.64 - 1.17	1.29	0.86 - 1.93
p-value (global)					< 0.01		0.46

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

6.3 Coffee

6.3.1 Analysis of the frequency and average daily amount of consumption

Overall, coffee frequency of intake was associated with a reduced risk of lung cancer, although regular intake was not statistically significantly associated in the adjusted

frequency of consumption model (table 10). The global test was statistically significant in the adjusted model ($p < 0.01$), confirming the significance of the overall association between the frequency of coffee consumption and lung cancer.

Table 10: OR estimates for the frequency of coffee consumption in relation to lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI
Never	89	88	1.00	Ref	1.00	Ref
Occasional	32	108	0.29	0.18 - 0.48	0.32	0.17 - 0.59
Regular	888	1238	0.71	0.52 - 0.96	0.77	0.52 - 1.16
p-value (global)				< 0.01		< 0.01

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

A similar pattern was observed when analyzing the average daily amount of consumption (table 11), although intake of > 5 cups/day was associated with a 65% increased risk of lung cancer in the unadjusted model, which became null in the multivariate model. While it appeared that a linear dose response was present, the trend test was not statistically significant in the adjusted model ($p = 0.10$).

Table 11: OR estimates for the average daily amount of coffee consumption in relation to Lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI
Never	89	88	1.00	ref	1.00	ref
<1 cup/day	33	114	0.29	0.18 - 0.47	0.32	0.17 - 0.59
1 cup/day	202	414	0.48	0.34 - 0.68	0.73	0.47 - 1.12
>1 to 3 cups/day	411	616	0.66	0.48 - 0.91	0.75	0.49 - 1.13
>3 to 5 cups/day	159	133	1.18	0.81 - 1.72	0.86	0.53 - 1.40
>5 cups/day	115	69	1.65	1.08 - 2.51	1.03	0.60 - 1.77
p-value (trend)				< 0.01		0.10
p-value (global)				< 0.01		< 0.01

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

6.3.2 Analysis of the duration of consumption

A consumption of coffee of ≥ 50 years was statistically significantly associated with a reduction of 36% of the risk of lung cancer in the unadjusted model. However, only statistically non-significant inverse associations were observed in the adjusted model (table 12). No trend was observed in either model.

Table 12: OR estimates for the duration of coffee consumption in relation to lung cancer

			Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI
Never	89	88	1.00	ref	1.00	ref
< 20 years	44	68	0.64	0.40 - 1.03	0.76	0.41 - 1.41
≥ 20 to < 30 years	99	134	0.73	0.49 - 1.08	0.94	0.55 - 1.61
≥ 30 to < 40 years	223	310	0.71	0.51 - 1.00	0.71	0.45 - 1.12
≥ 40 to < 50 years	347	455	0.75	0.54 - 1.05	0.80	0.52 - 1.23
≥ 50 years	175	271	0.64	0.45 - 0.91	0.75	0.46 - 1.21
p-value (trend)				0.09		0.27
p-value (global)				0.22		0.62

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income , and years of education

6.3.3 Analysis of the cumulative consumption

When examining the cumulative consumption in the unadjusted model (table 13), statistically significant inverse associations were observed for all subjects consuming ≤ 1 cup/day, but for subjects consuming > 1 cup/day, only non-significant inverse associations were observed. Although only non-significant inverse associations were observed in the adjusted model, there was a tendency for the inverse associations to be stronger in subjects consuming ≤ 1 cup/day.

Table 13: OR estimates for the cumulative consumption of coffee in relation to lung cancer

				Crude		Adjusted*	
		Cases	Controls	OR	95% CI	OR	95% CI
Never	-	89	88	1.00	ref	1.00	ref
≤ 1 cup/day	< 30 years	29	64	0.45	0.26 - 0.76	0.70	0.36 - 1.36
	≥ 30 to < 50 years	124	246	0.50	0.35 - 0.72	0.72	0.45 - 1.16
	≥ 50 years	50	110	0.45	0.29 - 0.70	0.71	0.40 - 1.27
> 1 cup/day	< 30 years	114	138	0.82	0.56 - 1.20	0.99	0.59 - 1.67
	≥ 30 to < 50 years	446	519	0.85	0.62 - 1.17	0.78	0.51 - 1.18
	≥ 50 years	125	161	0.77	0.53 - 1.12	0.74	0.44 - 1.23
p-value (global)					< 0.01		0.68

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income , and years of education

6.4 Secondary analysis

6.4.1 Black tea

6.4.1.1 Stratification by sex

When stratifying by sex, black tea consumption was generally slightly more strongly inversely associated with lung cancer among females then males for each of the various representations of tea intake (tables 14-17). It is important to note however that the p-values for interaction by sex (from the adjusted model) were non-significant for all analyses, ranging from 0.21 to 0.50. Moreover, in some cases, a stronger positive association was observed in female subjects when compared to male subjects, although the results observed were too imprecise to draw conclusions.

Table 14: OR estimates for the frequency of black tea consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	356	420	1.00	ref	1.00	ref	195	184	1.00	ref	1.00	ref	
Occasional	69	132	0.62	0.45 - 0.85	0.90	0.61 - 1.33	58	149	0.37	0.26 - 0.53	0.69	0.42 - 1.14	
Regular	240	319	0.89	0.71 - 1.11	1.14	0.87 - 1.49	193	265	0.69	0.52 - 0.90	0.98	0.66 - 1.44	
p-value (global)				0.01		0.44				< 0.01		0.31	
													0.43

* adjusted for age, sex, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 15: OR estimates for the average daily amount of black tea consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	356	420	1.00	ref	1.00	ref	195	184	1.00	ref	1.00	ref	
< 1	69	132	0.62	0.45 - 0.85	0.90	0.61 - 1.33	58	156	0.35	0.24 - 0.50	0.66	0.40 - 1.09	
1	67	129	0.61	0.44 - 0.85	0.87	0.59 - 1.28	61	102	0.56	0.39 - 0.82	1.29	0.76 - 2.18	
>1 to 3	133	155	1.01	0.77 - 1.33	1.33	0.96 - 1.85	96	131	0.69	0.50 - 0.96	0.87	0.54 - 1.39	
>3	40	35	1.35	0.84 - 2.17	1.23	0.70 - 2.16	36	25	1.36	0.79 - 2.35	0.88	0.42 - 1.83	
p-value (trend)				0.96		0.16				0.51		0.87	
p-value (global)				<0.01		0.25				<0.01		0.28	
													0.21

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 16: OR estimates for the duration of black tea consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	356	420	1.00	ref	1.00	ref	195	184	1.00	ref	1.00	ref	
< 30	42	47	1.05	0.68 - 1.64	1.34	0.80 - 2.26	49	69	0.67	0.44 - 1.02	1.24	0.70- 2.22	
≥ 30 to < 40	44	49	1.06	0.69 - 1.63	1.31	0.78 - 2.20	36	57	0.60	0.38 - 0.95	0.67	0.36 - 1.25	
≥40 to < 50	84	128	0.77	0.57 - 1.06	0.94	0.65 - 1.36	57	75	0.72	0.48 - 1.07	0.89	0.49 - 1.59	
≥ 50	70	70	0.87	0.62 - 1.22	1.27	0.82 - 1.95	51	64	0.75	0.49 - 1.14	1.33	0.69 - 2.57	
p-value (trend)				0.16		0.45				0.06		0.95	
p-value (global)				0.48		0.54				0.09		0.41	
													0.50

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 17: OR estimates for the cumulative consumption of black tea in relation to lung cancer stratified by sex

		Males						Females						
				Crude		Adjusted*				Crude		Adjusted*		
Cups/ day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	-	356	420	1.00	ref	1.00	ref	195	184	1.00	ref	1.00	ref	
≤ 1	< 30	15	22	0.80	0.41-1.57	0.94	0.41-2.13	21	34	0.58	0.33-1.04	1.95	0.87-4.35	
	≥ 30 to < 50	31	76	0.48	0.31-0.75	0.69	0.42-1.15	25	55	0.43	0.26-0.72	0.93	0.46-1.91	
	≥ 50	21	31	0.80	0.45-1.42	1.38	0.67-2.82	15	20	0.71	0.35-1.42	1.04	0.39-2.77	
> 1	< 30	27	25	1.27	0.73-2.24	1.62	0.85-3.09	28	35	0.76	0.44-1.29	0.90	0.44-1.85	
	≥ 30 to < 50	97	101	1.13	0.83-1.55	1.31	0.90-1.91	68	77	0.83	0.57-1.22	0.71	0.41-1.22	
	≥ 50	49	64	0.90	0.61-1.35	1.22	0.75-2.00	36	44	0.77	0.48-1.25	1.46	0.68-3.12	
p-value (global)					0.03		0.24				0.05		0.37	
														0.37

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

6.4.1.2 Stratification by smoking

When examining light and heavy smokers separately, no significant differences were observed in the association between the frequency of black tea consumption and lung cancer (table 18). Moreover, the interaction by smoking was statistically non-significant (p = 0.83). While more variations were observed between light and heavy smokers in the average daily amount of consumption analysis (table 19), the p-value for the interaction was statistically non-significant (p = 0.38). There were some indications that a shorter duration of consumption (table 20) was more strongly positively associated with lung cancer for heavy smokers, while a longer duration of consumption was more strongly positively associated with lung cancer among light smokers. However, the interaction by smoking was again statistically non-significant (p = 0.11). This tendency was also observed in the cumulative consumption analysis (table 21), but only for subjects consuming > 1 cup/day. For subjects consuming ≤ 1 cup/day, weaker positive associations and stronger inverse associations were generally observed in heavy smokers. The interaction by smoking was non-significant (p = 0.60). Overall, we observed no evidence that the association between black tea consumption and lung cancer vary by smoking level.

Table 18: OR estimates for the frequency of black tea consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	171	455	1.00	ref	1.00	ref	380	149	1.00	ref	1.00	ref	
Occasional	53	240	0.59	0.42 - 0.83	0.82	0.55 - 1.23	74	41	0.71	0.46 - 1.08	0.82	0.51 - 1.33	
Regular	159	465	0.91	0.71 - 1.17	1.12	0.83 - 1.51	274	119	0.90	0.68 - 1.20	1.03	0.75 - 1.43	
p-value global				0.01		0.32				0.27		0.65	
													0.83

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 19: OR estimates for the average daily amount of black tea consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	171	455	1.00	ref	1.00	ref	380	149	1.00	ref	1.00	ref	
< 1	53	247	0.57	0.40 - 0.81	0.79	0.53 - 1.18	74	41	0.71	0.46 - 1.08	0.83	0.51 - 1.33	
1	57	186	0.82	0.58 - 1.15	1.23	0.82 - 1.84	71	45	0.62	0.41 - 0.94	0.77	0.49 - 1.23	
>1 to 3	81	232	0.93	0.68 - 1.26	1.05	0.73 - 1.51	148	54	1.08	0.75 - 1.55	1.28	0.86 - 1.93	
>3	21	40	1.40	0.80 - 2.44	1.26	0.66 - 2.42	55	20	1.08	0.63 - 1.86	0.98	0.54 - 1.78	
p-value (trend)				0.81		0.39				0.99		0.57	
p-value (global)				<0.01		0.44				0.08		0.36	
													0.38

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 20: OR estimates for the duration of black tea consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	171	455	1.00	ref	1.00	ref	380	149	1.00	ref	1.00	ref	
< 30	38	98	1.03	0.68 - 1.56	1.24	0.76 - 2.02	53	18	1.16	0.66 - 2.04	1.31	0.71 - 2.42	
≥ 30 to < 40	21	87	0.64	0.39 - 1.07	0.71	0.39- 1.28	59	19	1.22	0.70 - 2.11	1.28	0.70 - 2.32	
≥40 to < 50	50	152	0.88	0.61 - 1.26	1.14	0.74 - 1.75	91	51	0.70	0.47 - 1.04	0.77	0.49 - 1.19	
≥ 50	50	128	1.04	0.72 - 1.51	1.46	0.91 - 2.36	71	31	0.90	0.57 - 1.43	1.08	0.63 - 1.86	
p-value (trend)				0.65		0.25				0.23		0.79	
p-value (global)				0.46		0.27				0.31		0.47	
													0.11

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 21: OR estimates for the cumulative consumption of black tea in relation to lung cancer stratified by smoking

		Light smokers						Heavy smokers						
				Crude		Adjusted*				Crude		Adjusted*		
Cups/ day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	-	171	455	1.00	ref	1.00	ref	380	149	1.00	ref	1.00	ref	
≤ 1	< 30	18	49	0.98	0.55-1.73	1.47	0.75-2.89	18	7	1.01	0.41-2.46	1.07	0.41-2.83	
	≥ 30 to < 50	22	102	0.57	0.35-0.94	0.94	0.54-1.64	34	29	0.46	0.27-0.78	0.58	0.33-1.04	
	≥ 50	17	42	1.08	0.60-1.94	1.42	0.69-2.94	19	9	0.83	0.37-1.87	1.15	0.47-2.88	
> 1	< 30	20	49	1.09	0.63-1.88	1.10	0.58-2.06	35	11	1.25	0.62-2.52	1.43	0.68-3.03	
	≥30 to < 50	49	137	0.95	0.66-1.38	1.00	0.64-1.53	116	41	1.11	0.74-1.66	1.14	0.73-1.79	
	≥ 50	33	86	1.02	0.66-1.58	1.47	0.83-2.45	52	22	0.93	0.54-1.58	1.12	0.61-2.05	
p-value (global)					0.49		0.70				0.12		0.48	
														0.60

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

6.4.1.3 Analysis by histological type

The association between the frequency of consumption (table 22) and the average daily amount of consumption (table 23) of black tea and lung cancer was similar for squamous cell carcinoma cases and adenocarcinoma cases, however, the unadjusted estimates suggested a positive association between the consumption of > 3 cups/day and squamous cell carcinoma. For small cell carcinoma cases, slightly stronger inverse associations and more null positive associations were generally observed. Similar results were observed in the duration of consumption analysis (table 24), but the adjusted estimates suggested a statistically significant positive association between a duration of consumption of ≥ 50 years and adenocarcinoma (OR: 1.63 (1.03 - 2.57)). A similar positive association was observed for adenocarcinoma cases in the cumulative consumption analysis (table 25). However, the association was only borderline statistically significant for adenocarcinoma cases consuming > 1 cup/day (OR: 1.68 (1.00 - 2.83)). This could indicate that a high

consumption of black tea over a long time period increase the risk of developing adenocarcinoma. However, since the global test for the adjusted duration model ($p = 0.20$), and for the adjusted cumulative consumption model ($p = 0.14$) were statistically non-significant, it is possible that the observed associations are the result of chance

Table 22: OR estimates for the frequency of black tea consumption in relation to lung cancer by histology

	Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
			Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	153	604	1.00	ref	1.00	ref	208	604	1.00	ref	1.00	ref	100	604	1.00	ref	1.00	ref
Occasional	35	281	0.49	0.33 - 0.73	0.84	0.53 - 1.34	54	281	0.56	0.40 - 0.78	0.85	0.57 - 1.26	19	281	0.41	0.25 - 0.68	0.78	0.42 - 1.44
Regular	134	584	0.91	0.70 - 1.17	1.11	0.81 - 1.53	176	584	0.88	0.70 - 1.10	1.15	0.86 - 1.53	64	584	0.66	0.47 - 0.92	0.85	0.56- 1.31
p-value (global)				<0.01		0.48				<0.01		0.31				<0.01		0.64

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 23: OR estimates for the daily amount of black tea consumption in relation to lung cancer by histology

	Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
			Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	153	604	1.00	ref	1.00	ref	208	604	1.00	ref	1.00	ref	100	604	1.00	ref	1.00	ref
< 1	35	288	0.48	0.32 - 0.71	0.83	0.52 - 1.32	54	288	0.54	0.39 - 0.76	0.83	0.56 - 1.23	19	288	0.40	0.24 - 0.66	0.76	0.41 - 1.41
1	36	231	0.62	0.42 - 0.91	0.82	0.51 - 1.30	54	231	0.68	0.49 - 0.95	1.04	0.69 - 1.54	22	231	0.58	0.35 - 0.94	0.83	0.46 - 1.52
>1 to 3	68	286	0.94	0.68 - 1.29	1.21	0.83 - 1.79	98	286	1.00	0.75 - 1.31	1.28	0.91 - 1.81	33	286	0.70	0.46 - 1.06	0.83	0.49- 1.42
>3	30	60	1.97	1.23 - 3.17	1.61	0.91 - 2.83	24	60	1.16	0.71 - 1.91	1.06	0.58 - 1.91	9	60	0.91	0.44 - 1.88	1.07	0.46 - 2.50
p-value (trend)				0.32		0.15	0.26			0.97		0.22				0.08		0.63
p-value (global)				<0.01		0.19	0.39			<0.01		0.40				<0.01		0.87

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 24: OR estimates for the duration of black tea consumption in relation to lung cancer by histology

	Squamous cell carcinoma		Adenocarcinoma		Small cell carcinoma	
	Crude	Adjusted*	Crude	Adjusted*	Crude	Adjusted*

Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	153	604	1.00	ref	1.00	ref	208	604	1.00	ref	1.00	ref	100	604	1.00	ref	1.00	ref
< 30	25	116	0.85	0.53 - 1.36	1.37	0.79- 2.38	39	116	0.98	0.66 - 1.45	1.17	0.72 - 1.89	14	116	0.73	0.40 - 1.32	1.12	0.55 - 2.30
≥ 30 to < 40	24	106	0.89	0.56 - 1.44	1.20	0.68 - 2.11	35	106	0.96	0.63 - 1.45	1.13	0.68 - 1.86	10	106	0.57	0.29 - 1.13	0.72	0.32 - 1.63
≥ 40 to < 50	48	203	0.93	0.65 - 1.34	1.03	0.67 - 1.58	49	203	0.70	0.49 - 0.99	0.90	0.59 - 1.37	23	203	0.68	0.42 - 1.11	0.71	0.39 - 1.30
≥ 50	37	159	0.92	0.62 - 1.37	1.05	0.63 - 1.76	53	159	0.97	0.68 - 1.37	1.63	1.03 - 2.57	17	159	0.65	0.38 - 1.11	0.94	0.46 - 1.94
p-value (trend)				0.61		0.79				0.24		0.21				0.03		0.38
p-value (global)				0.95		0.83				0.40		0.20				0.19		0.74

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 25: OR estimates for the cumulative consumption of black tea in relation to lung cancer by histology

		Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
				Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
Cups/day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	-	153	604	1.00	ref	1.00	ref	208	604	1.00	ref	1.00	ref	100	604	1.00	ref	1.00	ref
≤ 1	< 30	6	56	0.42	0.18 - 1.00	0.74	0.27 - 2.00	18	56	0.93	0.54 - 1.62	1.47	0.75 - 2.87	8	56	0.86	0.40 - 1.86	1.35	0.52 - 3.52
	≥ 30 to < 50	17	131	0.51	0.30 - 0.88	0.64	0.35 - 1.19	20	131	0.44	0.27 - 0.73	0.65	0.37 - 1.15	8	131	0.37	0.18 - 0.78	0.53	0.22 - 1.28
	≥ 50	13	51	1.01	0.53 - 1.90	1.36	0.62 - 3.03	16	51	0.91	0.51 - 1.63	1.57	0.76 - 3.22	6	51	0.71	0.30 - 1.70	0.93	0.30 - 2.88
> 1	< 30	19	60	1.25	0.72 - 2.16	1.80	0.96 - 3.40	21	60	1.02	0.60 - 1.71	0.97	0.52 - 1.79	6	60	0.60	0.25 - 1.44	0.93	0.35 - 2.50
	≥ 30 to < 50	55	178	1.22	0.86 - 1.73	1.38	0.90 - 2.10	64	178	1.04	0.75 - 1.45	1.19	0.79 - 1.77	25	178	0.85	0.53 - 1.36	0.81	0.45 - 1.46
	≥ 50	24	108	0.88	0.55 - 1.41	0.96	0.53 - 1.74	37	108	1.00	0.66 - 1.49	1.68	1.00 - 2.83	11	108	0.62	0.32 - 1.19	0.94	0.40 - 2.21
p-value (global)					0.04		0.15				0.09		0.14				0.16		0.83

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

6.4.2 Coffee

6.4.2.1 Stratification by sex

When stratifying the frequency of consumption and the average daily amount of consumption of coffee by sex (tables 26-27), similar associations were observed for males and females in the unadjusted models. For the adjusted models, however, the inverse associations were generally stronger and only significant in males. The association between the duration of coffee consumption (table 28) and lung cancer was similar between males and females in the unadjusted model. In the adjusted model, however, while only statistically non-significant inverse associations were observed for males, statistically non-significant positive associations were observed at shorter durations of consumption for females. Similar observations were made in the cumulative consumption analysis (table 29). The interaction by sex was non-significant in all models, with p-values ranging from 0.52 to 0.98. Overall, we did not observe any indication that the association between coffee consumption and lung cancer statistically significantly varied by sex.

Table 26: OR estimates for the frequency of coffee consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	53	53	1.00	ref	1.00	ref	36	35	1.00	ref	1.00	ref	
Occasional	18	70	0.26	0.14 - 0.49	0.23	0.11 - 0.49	14	38	0.36	0.17 - 0.77	0.66	0.22 - 1.94	
Regular	507	717	0.71	0.48 - 1.05	0.69	0.42 - 1.14	381	521	0.71	0.44 - 1.15	0.99	0.49 - 2.00	
p-value (global)				< 0.01		< 0.01				< 0.03		0.66	
													0.30

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 27: OR estimates for the average daily amount of coffee consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	53	53	1.00	ref	1.00	ref	36	35	1.00	ref	1.00	ref	
≤ 1	19	73	0.26	0.14 - 0.49	0.24	0.11 - 0.51	14	41	0.33	0.16 - 0.71	0.58	0.20 - 1.66	
1	127	236	0.54	0.35 - 0.83	0.69	0.40 - 1.19	75	178	0.41	0.24 - 0.70	0.85	0.40 - 1.82	
> 1 to 3	235	350	0.67	0.44 - 1.02	0.68	0.40 - 1.14	176	266	0.64	0.39 - 1.06	0.94	0.46 - 1.94	
> 3 to 5	81	84	0.96	0.59 - 1.57	0.65	0.36 - 1.20	78	49	1.55	0.86 - 2.78	1.44	0.63 - 3.28	
> 5	63	44	1.43	0.83 - 2.46	0.87	0.45 - 1.68	52	25	2.02	1.04 - 3.94	1.48	0.57 - 3.81	
p-value (trend)				< 0.01		0.47				< 0.01		0.08	
p-value (global)				< 0.01		< 0.01				< 0.01		0.30	
													0.52

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 28: OR estimates for the duration of coffee consumption in relation to lung cancer stratified by sex

	Males						Females						
			Crude		Adjusted*				Crude		Adjusted*		
Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	53	53	1.00	ref	1.00	ref	36	35	1.00	ref	1.00	ref	
< 30	58	82	0.71	0.43 - 1.18	0.65	0.34 - 1.24	85	120	0.69	0.40 - 1.18	1.61	0.70 - 3.69	
≥ 30 to < 40	116	164	0.71	0.45 - 1.11	0.58	0.33 - 1.02	107	146	0.71	0.42 - 1.21	1.09	0.50 - 2.37	
≥ 40 to < 50	222	291	0.76	0.50 - 1.16	0.76	0.45 - 1.30	125	164	0.74	0.44 - 1.25	0.87	0.40 - 1.86	
≥ 50	111	180	0.62	0.39 - 0.97	0.74	0.41 - 1.34	64	91	0.68	0.39 - 1.20	0.76	0.32 - 1.80	
p-value (trend)				0.12		0.60				0.51		0.19	
p-value (global)				0.30		0.55				0.71		0.40	
													0.98

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

Table 29: OR estimates for the cumulative consumption of coffee in relation to lung cancer stratified by sex

		Males						Females						
				Crude		Adjusted*				Crude		Adjusted*		
Cups/ day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	-	53	53	1	ref	1	ref	36	35	1	ref	1	ref	
≤ 1	< 30	11	27	0.41	0.18-0.91	0.47	0.18-1.19	18	37	0.47	0.23-0.98	1.26	0.46-3.45	
	≥ 30 to < 50	82	141	0.58	0.36-0.93	0.74	0.41-1.32	42	105	0.39	0.22-0.70	0.73	0.32-1.68	
	≥ 50	35	71	0.49	0.28-0.86	0.69	0.35-1.38	15	39	0.37	0.18-0.80	0.81	0.29-2.33	
> 1	< 30	47	55	0.86	0.50-1.47	0.80	0.40-1.59	67	83	0.79	0.45-1.38	1.69	0.72-3.97	
	≥ 30 to < 50	256	314	0.82	0.54-1.24	0.67	0.40-1.13	190	205	0.90	0.54-1.49	1.07	0.51-2.25	
	≥ 50	76	109	0.70	0.43-1.13	0.71	0.38-1.32	49	52	0.92	0.50-1.68	0.78	0.31-1.97	
p-value (global)					0.03		0.73				< 0.01		0.48	
														0.87

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by sex in the adjusted model

6.4.2.2 Stratification by smoking

The association between the consumption of coffee and lung cancer (tables 30-33) was similar for light and heavy smokers. Slightly stronger inverse associations tended to be observed in light smokers, and in the adjusted model the global test was only statistically significant for light smokers. However, the interaction by smoking was never statistically significant (p-values ranging from 0.27 to 0.88) and we did not observe any indication that smoking modified the association between coffee consumption and lung cancer.

Table 30: OR estimates for the frequency of coffee consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	38	73	1.00	ref	1.00	ref	51	15	1.00	ref	1.00	ref	
Occasional	10	89	0.22	0.10 - 0.46	0.26	0.11 - 0.61	22	19	0.34	0.15 - 0.79	0.36	0.14 - 0.95	
Regular	312	978	0.61	0.41 - 0.93	0.68	0.41 - 1.12	576	260	0.65	0.36 - 1.18	0.85	0.44 - 1.66	
p-value (global)				< 0.01		< 0.01				0.04		0.06	
													0.88

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 31: OR estimates for the average daily amount of coffee consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	38	73	1.00	ref	1.00	ref	51	15	1.00	ref	1.00	ref	
< 1	11	94	0.23	0.11 - 0.47	0.27	0.12 - 0.62	22	20	0.32	0.14 - 0.75	0.35	0.14 - 0.91	
1	81	356	0.44	0.28 - 0.69	0.60	0.34 - 1.04	121	58	0.61	0.32 - 1.18	0.91	0.44 - 1.89	
> 1 to 3	154	490	0.60	0.39 - 0.93	0.66	0.39 - 1.12	257	126	0.60	0.33 - 1.11	0.83	0.42 - 1.64	
> 3 to 5	43	88	0.94	0.55 - 1.60	0.76	039 - 1.45	116	45	0.76	0.39 - 1.48	0.92	0.43 - 1.95	
> 5	33	39	1.63	0.89 - 2.98	1.39	0.66 - 2.93	82	30	0.80	0.40 - 1.64	0.77	0.35 - 1.72	
p-value (trend)				< 0.01		0.14				< 0.01		0.76	
p-value (global)				< 0.01		< 0.01				< 0.01		0.25	
													0.50

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 32: OR estimates for the duration of coffee consumption in relation to lung cancer stratified by smoking

	Light smokers						Heavy smokers						
			Crude		Adjusted*				Crude		Adjusted*		
Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	38	73	1.00	ref	1.00	ref	51	15	1.00	ref	1.00	ref	
< 30	82	180	0.88	0.55 - 1.40	0.80	0.43 - 1.49	61	22	0.82	0.38 - 1.73	0.93	0.39 - 2.19	
≥ 30 to < 40	63	237	0.51	0.32 - 0.83	0.56	0.31 - 1.01	160	73	0.65	0.34 - 1.22	0.77	0.37 - 1.60	
≥ 40 to < 50	103	353	0.56	0.36 - 0.88	0.65	0.37 - 1.13	244	102	0.70	0.38 - 1.31	0.96	0.48 - 1.92	
≥ 50	64	208	0.59	0.37 - 0.96	0.71	0.38 - 1.32	111	63	0.52	0.27 - 1.00	0.80	0.37 - 1.70	
p-value (trend)				< 0.01		0.19				0.05		0.69	
p-value (global)				< 0.01		0.28				0.25		0.80	
													0.32

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

Table 33: OR estimates cumulative consumption of coffee in relation to lung cancer stratified by smoking

		Light smokers						Heavy smokers						
				Crude		Adjusted*				Crude		Adjusted*		
Cups/ day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	P-int**
Never	-	38	73	1.00	ref	1.00	ref	51	15	1.00	ref	1.00	ref	
≤ 1	< 30	16	59	0.52	0.27-1.03	0.56	0.25-1.26	13	5	0.77	0.24-2.49	1.12	0.31-3.98	
	≥ 30to < 50	46	210	0.42	0.25-0.70	0.62	0.34-1.13	78	36	0.64	0.32-1.28	0.83	0.38-1.79	
	≥ 50	20	92	0.42	0.22-0.78	0.52	0.25-1.11	30	18	0.49	0.22-1.11	1.01	0.40-2.53	
> 1	< 30	66	121	1.05	0.64-1.72	0.98	0.51-1.85	48	17	0.83	0.37-1.85	0.95	0.39-2.33	
	≥ 30 to < 50	120	380	0.61	0.39-0.94	0.61	0.35-1.04	326	139	0.69	0.38-1.27	0.90	0.46-1.77	
	≥ 50	44	116	0.73	0.43-1.23	0.80	0.41-1.55	81	45	0.53	0.27-1.05	0.66	0.30-1.46	
p-value (global)					< 0.01		0.20				0.50		0.91	
														0.27

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education
** P-value for the interaction by smoking in the adjusted model

6.4.2.3 Analysis by histological type

The association between the frequency of coffee consumption and lung cancer (table 34) was similar for squamous cell carcinoma cases and adenocarcinoma cases, but inverse associations were generally stronger and positive associations closer to the null value for small cell carcinoma cases. The adjusted estimates were never statistically significant for adenocarcinoma cases and the global test was statistically significant only for small cell carcinoma cases ($p = 0.03$). Similar results were observed in the average daily amount of consumption, duration of consumption and cumulative consumption analyses (tables 34-36). Overall, there is some evidence that coffee is more strongly associated with small cell carcinoma; however, due to the low statistical power of our analyses, the evidence is weak.

Table 34: OR estimates for the frequency of coffee consumption in relation to lung cancer by histology

	Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
			Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	24	88	1.00	ref	1.00	ref	29	88	1.00	ref	1	ref	22	88	1.00	ref	1.00	ref
Occasional	9	108	0.31	0.14 - 0.69	0.37	0.14 - 0.94	17	108	0.48	0.25 - 0.93	0.53	0.24 - 1.16	2	108	0.07	0.02 - 0.32	0.11	0.02 - 0.56
Regular	254	1238	0.75	0.47 - 1.21	0.90	0.50 - 1.62	356	1238	0.87	0.46 - 1.35	0.97	0.56 - 1.67	141	1238	0.46	0.28 - 0.75	0.66	0.32 - 1.34
p-value (global)				0.02		0.07				0.06		0.15				< 0.01		0.03

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 35: OR estimates for the average daily amount of coffee consumption in relation to lung cancer by histology

	Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
			Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
Cups/day	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	24	88	1.00	ref	1.00	ref	29	88	1.00	ref	1	ref	22	88	1.00	ref	1.00	ref
< 1	9	114	0.29	0.13 - 0.65	0.36	0.14 - 0.91	18	114	0.48	0.25 - 0.92	0.54	0.25 - 1.18	2	114	0.07	0.02 - 0.31	0.10	0.02 - 0.53
1	57	414	0.51	0.30 - 0.86	0.81	0.43 - 1.55	76	414	0.56	0.34 - 0.91	0.85	0.47 - 1.53	31	414	0.30	0.17 - 0.54	0.67	0.30 - 1.48
>1 to 3	122	616	0.73	0.44 - 1.19	0.88	0.48 - 1.62	174	616	0.86	0.55 - 1.35	0.99	0.57 - 1.74	61	616	0.40	0.23 - 0.68	0.57	0.28 - 1.20
>3 to 5	49	133	1.35	0.77 - 2.36	1.14	0.57 - 2.28	59	133	1.35	0.80 - 2.26	1.06	0.56 - 2.01	23	133	0.69	0.36 - 1.32	0.62	0.26 - 1.48
>5	26	69	1.38	0.73 - 2.62	1.02	0.46 - 2.23	46	69	2.02	1.15 - 3.55	1.16	0.58 - 2.32	26	69	1.51	0.79 - 2.89	1.36	0.56 - 3.31
p-value (trend)				< 0.01		0.14				< 0.01		0.14				< 0.01		0.22
p-value (global)				< 0.01		0.16				< 0.01		0.39				< 0.01		0.01

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 36: OR estimates for the duration of coffee consumption in relation to lung cancer by histology

	Squamous cell carcinoma	Adenocarcinoma	Small cell carcinoma

			Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	24	88	1.00	ref	1.00	ref	29	88	1.00	ref	1	ref	22	88	1.00	ref	1.00	ref
< 30	30	202	0.55	0.30 - 0.99	0.91	0.43 - 1.95	68	202	1.02	0.62 - 1.69	1.20	0.62 - 2.30	16	202	0.32	0.16 - 0.63	0.52	0.20 - 1.35
≥ 30 to < 40	61	310	0.72	0.43 - 1.22	1.02	0.52 - 1.99	94	310	0.92	0.57 - 1.49	0.93	0.51 - 1.70	41	310	0.53	0.30 - 0.94	0.61	0.27 - 1.36
≥ 40 to < 50	108	455	0.87	0.53 - 1.43	0.97	0.52 - 1.82	129	455	0.86	0.54 - 1.37	0.92	0.51 - 1.65	58	455	0.51	0.30 - 0.88	0.72	0.34 - 1.54
≥ 50	55	271	0.74	0.44 - 1.27	0.68	0.34 - 1.36	65	271	0.73	0.44 - 1.20	0.95	0.50 - 1.82	26	271	0.38	0.21 - 0.71	0.72	0.29 - 1.75
p-value (trend)				0.63		0.39				0.07		0.58				0.10		0.67
p-value (global)				0.20		0.57				0.47		0.84				< 0.01		0.73

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

Table 37: OR estimates for the cumulative consumption of coffee in relation to lung cancer by histology

		Squamous cell carcinoma						Adenocarcinoma						Small cell carcinoma					
				Crude		Adjusted*				Crude		Adjusted*				Crude		Adjusted*	
Cups/day	Years	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI	Cases	Controls	OR	95% CI	OR	95% CI
Never	-	24	88	1.00	ref	1.00	ref	29	88	1.00	ref	1.00	ref	22	88	1.00	ref	1.00	ref
	< 30	6	64	0.34	0.13 - 0.89	0.77	0.26 - 2.28	14	64	0.66	0.33 - 1.36	1.09	0.47 - 2.51	4	64	0.25	0.08 - 0.76	0.56	0.15 - 2.11
≤ 1	≥ 30 to < 50	39	246	0.58	0.33 - 1.02	0.95	0.48 - 1.89	39	246	0.48	0.28 - 0.82	0.64	0.33 - 1.24	22	246	0.36	0.19 - 0.68	0.71	0.30 - 1.67
	≥ 50	12	110	0.40	0.19 - 0.85	0.52	0.21 - 1.26	24	110	0.66	0.36 - 1.22	1.18	0.56 - 2.50	5	110	0.18	0.07 - 0.50	0.57	0.17 - 1.92
	< 30	24	138	0.64	0.34 - 1.19	0.96	0.43 - 2.13	54	138	1.19	0.70 - 2.01	1.28	0.65 - 2.51	12	138	0.35	0.16 - 0.74	0.54	0.20 - 1.48
> 1	≥ 30 to < 50	130	519	0.92	0.56 - 1.50	1.02	0.55 - 1.88	184	519	1.08	0.69 - 1.69	1.03	0.59- 1.81	77	519	0.59	0.35 - 1.00	0.66	0.31 - 1.39
	≥ 50	43	161	0.98	0.56 - 1.72	0.78	0.37 - 1.62	41	161	0.77	0.45 - 1.33	0.82	0.41 - 1.64	21	161	0.52	0.27 - 1.00	0.74	0.29 - 1.88
p-value (global)					< 0.01		0.68				< 0.01		0.33				< 0.01		0.92

* adjusted for age, smoking, wine intake, energy intake, proxy status, ethnicity, income, and years of education

7.0 Discussion

7.1 Summary of key finding

7.1.1 Black tea

The role of black tea in the development of lung cancer is hard to predict based on what is known of its chemical composition. While similar to green tea, the oxidative fermentation procedure required for the production of black tea reduces the concentration and bioavailability of antioxidants in the beverage (12). Moreover, little is known regarding how all of the components composing the beverage interact with each other. While we observed no statistically significant association between the consumption of black tea and lung cancer, the occasional consumption of black tea was inversely associated with lung cancer, while a regular consumption was positively associated. Although stronger associations were observed at higher average daily amount of black tea consumption, no trend was observed. There was no indication that the association was modified by sex or smoking level and while a stronger inverse association was generally observed in small cell carcinoma cases, the lower statistical power of the analyses made it impossible to reach any clear conclusion.

More interesting results were observed in the duration of black tea consumption analysis. While no statistically significant association was observed, the association was non-monotonic. Moreover, although there was no indication that the association was modified by sex or smoking, a statistically significant positive association was observed for adenocarcinoma cases consuming black tea for 50 years or more. While the role of chance in those results is of concern, they indicate that the consumption of black tea over a long period of time could increase the risk of adenocarcinoma.

7.1.2 Coffee

Coffee, like black tea, is a complex beverage with a still unknown role in the development of lung cancer. However, contrary to tea, its consumption has historically not been well regarded, due in part to its association with unhealthy life habits such as smoking. Our observation of a strong and statistically significant inverse association for occasional coffee drinking and of a statistically non-significant inverse association for regular coffee drinking is, however, not so unexpected. Coffee is composed of a number of antioxidants (38, 39) that could be the source of the observed inverse association and a statistically significant inverse association has been previously reported (70). There were indications of the presence of a dose-response, with an increase in the average amount of cups of coffee consumed daily reducing the strength of the inverse association until it became null, but no trend was observed. There was no indication that the association between coffee consumption and lung cancer was modified by sex or smoking. Stronger inverse associations were observed in small cell carcinoma cases, but the low number of subjects, particularly of occasional drinkers with either squamous cell carcinoma (9 cases) or small cell carcinoma (2 cases) reduced the precision of the analyses.

The duration of consumption did not affect the association between the consumption of coffee and lung cancer and the results observed in the cumulative consumption analysis were similar to what was observed in the average daily amount of consumption analysis. There was no indication that the association was modified by sex or smoking and while stronger inverse associations were observed in small cell carcinoma cases, the low statistical power of the analyses make it difficult to reach any clear conclusion.

7.1.3 Comparison with previous studies

7.1.3.1 Black tea

The literature on the association between the consumption of black tea and lung cancer has been inconsistent, both in the result reported and the methodology used. There has been, however, a tendency for inverse associations to be reported. The non-significant inverse association that we observed for occasional black tea drinking was akin to what was reported by studies which compared similar levels of consumption (44, 58-60, 62, 67). Where our results tended to differ from those studies was on the reversal of the association observed at higher average daily amount of black tea consumption. The lack of meaningful associations observed for subjects consuming 1 or more cup of black tea per day had been reported previously (60, 62), but there was a tendency for studies which reported an inverse association to observe a stronger inverse association at higher daily amount of black tea consumption (58, 61, 67). Still, a stronger positive association at higher daily amount of black tea consumption has also been reported by others (62, 64, 65). It is possible that at higher levels of consumption the protective effect of the antioxidants present in black tea become weaker than the cancer promoting effect of the caffeine present in black tea. The reduction in the strength of the inverse association with an increase in the average daily amount of cups of coffee consumed observed in our coffee analysis would also be consistent with this hypothesis. While there has been some indications in previous research that the association between black tea consumption and lung cancer is modified by sex (44, 52, 56) and smoking (44, 61, 67), the evidence was weak due to low statistical power and contradictory observations. While our statistical power was not optimal, it was generally stronger than in previously mentioned studies. We reported no evidence of a statistically significant interaction by sex or smoking between black tea consumption and lung cancer, although statistical power was low for these analyses. The association between black tea and lung cancer was not reported to vary between the main histological types of lung cancer in the literature (44, 59). However, we observed stronger inverse associations in small cell carcinoma cases compared to adenocarcinoma cases and large cell carcinoma

cases. Considering our low statistical power in the histological analyses and what was previously reported in the literature, it is probable that the variation we observed is the result of chance.

To our knowledge, this is the first study to analyse the association between the duration of black tea consumption and lung cancer. In studies that analyzed the association between tea consumption and ovarian cancer (97) and skin cancer (98), a longer duration of consumption was associated with a stronger inverse association. Although we reported a non-monotonic association, a statistically significant positive association was only observed for adenocarcinoma cases who consumed black tea for at least 50 years. This observation is contradictory to what was reported in ovarian cancer and skin cancer and indicate that the mechanism through which a longer duration of tea consumption is associated with lung cancer is different from the mechanism found in those two types of cancer. Those results also indicate that the association between tea and cancer is stronger at longer durations of consumption and show the importance of considering the duration of consumption when analyzing this beverage.

7.1.3.2 Coffee

While positive associations between the consumption of coffee and lung cancer have tended to be reported in the literature, it is not the first time that inverse associations have been observed (44, 52, 54, 58, 59, 63, 64, 70). It is however the first time that such a strong statistically significant inverse association has been reported. The strongest previously reported statistically significant reduction in the risk of lung cancer was of 59% (70), while statistically non-significant reductions in risk have generally ranged from 10% to 50% (44, 52, 54, 58, 59, 63, 64). It is possible that the weaker reduction in risk reported in the literature was because of the comparison of subjects consuming higher daily quantities of coffee. Indeed, results observed from cohort studies (52, 71) have indicated

that the consumption of a higher amount of cups of coffee per day was positively associated with lung cancer when compared to a lower amount. Moreover, we observed in our analysis that as the average daily amount of cups of coffee consumed increased, the inverse association became more null. It is also possible that residual confounding by smoking resulted in the underestimation of the strength of the inverse association in previous studies.

Some studies observed that the association between coffee consumption and lung cancer varied by sex (52, 54, 69), however, none provided strong evidence. Stensvold et al (69) only observed a statistically significant positive association in males but reported very similar ORs for both sexes, while Takezaki et al (54) reported both weaker positive associations and stronger inverse associations in females but overlapping confidence intervals. Khan et al (52) reported opposite associations in females and males, however, only 10 females were included in the analysis. Moreover, one study reported similar associations in males and females (44). While we also observed no evidence that the association between coffee consumption and lung cancer varied by sex, our statistical power was very low. For smoking, there was a tendency for studies that included only non-smokers to report inverse associations between the consumption of coffee and lung cancer (58, 64), while the one study that only included smokers (67) principally reported positive associations. Another study reported that the statistically significant positive association observed was principally present in current smokers (71). However, one study reported that the interaction by smoking was statistically non-significant (44). We reported no interaction by smoking in the association between coffee consumption and lung cancer. While our analysis could have suffered from a type II error due to a low statistical power, the lack of strong evidence in the literature combined with our observation suggests that smoking does not modify the association. In the literature, stronger inverse associations were generally observed for adenocarcinoma cases (54, 59), which is different from what we reported. However, as was the case in our study, these analyses suffered from low statistical power and one study reported similar associations among all histological types (44). It is probable

that the variation in the strength of the association observed between histological types was due to chance.

To our knowledge, this is the first study to analyze the association between the duration of coffee consumption and lung cancer. Our results, however, are similar to what was reported by a pooled analysis of case-control studies on the association between coffee and bladder cancer (99), and indicate that the association between the consumption of coffee and lung cancer may be constant over time.

7.2 Limitations and sources of bias

A number of limitations and potential sources of bias were present in this study and need to be considered during the interpretation of the results.

7.2.1 Precision

While this study included a high number of subjects compared to similar studies, precision could still be an issue. A majority of the subjects did not consume black tea and only a minority of the subjects who consumed it had an average daily amount of consumption of more than 1 cup per day, reducing the precision of the analysis at higher average daily amounts of consumption. For coffee, there were few subjects who either never consumed coffee or were occasional coffee drinkers, reducing the precision of all coffee analyses, but particularly of the frequency of consumption and the average daily amount of consumption analyses. During the stratified analyses, and particularly the analysis by histological type, the number of subjects present in each category was further reduced, in some case to as few as 2 subjects, resulting in a low precision.

7.2.2 Selection bias

The response rates of 84.2% for cases and 69.4% for controls observed in the MLCS were high, but the possibility that a difference, which could reduce the internal validity of the study, existed between respondents and non-respondents must be taken into consideration. Cases were lung cancer patients selected from hospitals and it is possible that respondents suffered from less severe or less advance type of lung cancer when compared to non-respondents. In this case, our results would be more representative of those less severe types of lung cancer. As for controls, respondents could have been more cautious about their health and have a healthier lifestyle when compared to the rest of the population. Both tea and coffee consumption have been associated with health awareness (100), however, coffee consumption has been associated with unhealthier behaviors while tea consumption has been associated with healthier behaviors. As a consequence, controls could have consumed less coffee and more black tea compared to the rest of the population. This would result in the underestimation of the inverse association observed for coffee and of the positive association observed for black tea.

7.2.3 Information bias

7.2.3.1 Recall bias

Although the data used in this analysis was reported by the subject or his/her proxy, the consumption of black tea or coffee is not commonly associated with lung cancer. As such, it is unlikely that cases (or their proxies) tried harder to remember or associated their illness with those beverages and this study should not suffer from a recall bias.

7.2.3.2 Interviewer bias

In any case-control study where the interviewer is not blinded to the status of case or control of the subject there is always the possibility that subjects would be interviewed

differently based on their illness status. For example, it is possible that the interviewer tried harder to gather information on exposure for cases than for controls, resulting in a differential classification error. However, it is unlikely that our study suffered from this classification error as neither black tea nor coffee consumption is commonly associated with lung cancer and as interviewers were trained to insure that cases and controls would be treated equally during the interview.

7.2.3.3 Other measurement error

Due to the nature of the questionnaire, mistakes could have been made during the assessment of the main exposures. The question used to differentiate between occasional and regular drinkers was imprecise, asking subjects if they drank black tea or coffee nearly every day. What is considered to be "nearly" every day can vary based on the personal opinion of each subject, and it is possible that subjects with the same frequency of consumption classified themselves differently. Occasional drinkers did not indicate their average daily amount of consumption and it was assumed that a lower frequency of consumption was indicative of a lower average daily amount of consumption. However, it is possible that a subject classified as an occasional drinker had a higher average daily amount of consumption than a regular drinker. Because of this, great care needs to be taken when interpreting the result observed in the "less than 1 cup per day" category of the average daily amount of consumption analysis. We had no information regarding the brewing time or the methodology used to prepare each cup of black tea or coffee. However, Hakim et al (77) reported that the brewing time, concentration and temperature of a cup of black tea influenced the association between black tea intake and squamous cell carcinoma of the skin. Our lack of data regarding the methodology used to prepare a cup of black tea or coffee mean that subjects reporting lower consumptions of one of those beverages could have ingested higher concentration of it than subjects reporting higher consumption and vice versa. All of those measurement error could have resulted in the misclassification of subjects and in the dilution of the association between black tea or coffee consumption and

lung cancer observed in the frequency of consumption, average daily amount of consumption and cumulative consumption analyses. We assumed that our subjects consumed principally black tea, however, it is possible that the consumption of green tea or decaffeinated tea was higher than expected in our study population. Considering the literature on green tea (44, 48-56, 68) and the possible role of caffeine in the development of lung cancer (13), a higher than expected consumption of those types of tea would results in the underestimation of the positive association between black tea consumption and lung cancer. We did not differentiate between the various types of coffee in our study and it was assumed that our subjects consumed principally caffeinated Arabica or Robusta coffee. However, it is possible that our subjects consumed coffee with low or no caffeine content (decaffeinated coffee, espresso, cappuccino). This could result in the overestimation of the inverse association observed between coffee consumption and lung cancer.

The duration variable did not consider the time since the end of the consumption for subjects who no longer consumed the beverage. If a subject had stopped his consumption less than 2 years before the interview he was still considered as a current drinker. However, a number of subjects had stopped their consumption earlier and some could have stopped their consumption prior to the etiologically relevant period of time. As a consequence, some duration intervals could be less relevant than others and could potentially dilute the association in the duration of consumption and cumulative consumption analyses. This did not seems to be an issue, however, as excluding from the analysis the 114 subjects (69 cases and 45 controls) who were past consumers of black tea, and the 79 subjects (56 cases and 23 controls) who were past consumers of coffee did not significantly influenced the observed associations.

7.2.4 Confounding

We adjusted the models for a number of covariates but due to a lack of information we were not able to adjust for all potential covariates, which could have biased our results.

For example, we did not have information regarding the addition of milk or cream in each cup of black tea or coffee. Milk and milk products have been positively associated with lung cancer (63, 64) and are potentially positively associated with black tea and coffee consumption. Not including those potential covariates in our analyses could have resulted in a more null association for coffee and a stronger positive association for black tea. As a consequence, the associations we reported should not be considered as representative of the effect of coffee or black tea alone. Rather, they should be considered as representative of the cumulative effect of a mixture of those beverages and other elements generally added to them in North-America. Moreover, the backward deletion procedure was used to determine the addition of certain covariates in the model. Because covariates were only tested using the frequency of consumption variable, it is possible that for the average daily amount of consumption, the duration of consumption and the cumulative consumption models the adjustment was not optimal. Residual confounding by smoking could also still be an issue. Although we used the CSI to adjust for smoking, it is possible that the data used to build the index was inaccurate, reducing the validity of the CSI measurement.

7.3 Implication in public health

7.3.1 Black tea

We reported that the occasional consumption of black tea was inversely associated with lung cancer, while the regular consumption of black tea was weakly positively associated with lung cancer. Because none of those associations were statistically significant, we cannot reach any strong conclusion regarding the impact of the consumption of black tea in health based on the level of consumption alone. However, we reported that consuming black tea for 50 years or more was statistically significantly positively associated with adenocarcinoma. This observation is worrisome as black tea is one of the most consumed beverages worldwide, and it is not unusual for individuals to consume black tea over a majority of their lifetime. However, our results only provided weak

evidence that the long-term consumption of black tea was positively associated with lung cancer. They, however, showed the importance of considering the duration of consumption when analyzing black tea. To better understand the role of black tea in the development of lung cancer, further studies with repeated measurements of a well defined consumption of black tea over a long time period are required.

7.3.1 Coffee

The perception of the role of coffee in the development of lung cancer has been mixed, with inconsistent results observed in the literature. However, there was a tendency in the literature for coffee to be positively associated with lung cancer. This tendency could be a result of the difficulties associated with separating the effect of the consumption of coffee from other risk factors correlated to its consumption such as smoking (83-85). In our study we observed results that went against this tendency and which brought to a better light the possible role of coffee in health. Moreover, our observations were consistent with the statistically significant inverse association reported in a recently published large prospective cohort study (101), between the consumption of coffee and death from all causes. Our study did, however, suffer from a number of weaknesses that could have biased the results. Considering both the potential protective effect of a moderate coffee consumption and the worldwide ubiquity of its consumption, it would be important for our results to be reproduced in studies that would obtain repeated and detailed measurements of the consumption of coffee, consider the methodology used in the preparation of the beverage and take great care to include every potential covariates associated with its consumption.

8.0 Conclusion

We observed no statistically significant associations between the frequency of black tea consumption or the average daily amount of black tea consumption and lung cancer. There was a weak indication that consuming black tea for 50 years or more was positively associated with adenocarcinoma. There was no evidence that the association was modified by sex or smoking. For coffee, we observed a statistically significant inverse association between the occasional consumption of coffee and lung cancer. For a regular consumption, however, the association was weaker and statistically non-significant. While the strength of the inverse association became weaker as the average daily amount of cups of coffee consumed increased, no linear trend was observed. The duration of coffee consumption was not associated with lung cancer, indicating that the effect of the consumption of coffee may be constant over time. There was no evidence that the association was modified by sex or smoking and only minimal variation was observed between the histological types. Further analysis with consideration of the methodology used in the preparation of black tea and coffee, and with repeated measurements of the exposure over a long time period would be required to better define the role of those two beverages in the development of lung cancer.

References

1. IARC. Globoscan 2008. 2010. Retrieved from http://globocan.iarc.fr/. Assessed on January 2012.

2. Wu-William AH, Samet JM. Chapter 4: Lung cancer and cigarette smoking:71-101. Samet JM. Epidemiology of Lung Cancer Vol.74: epidemiology of Lung Cancer. New York: Marcel Dekker; 1994:

3. Spitz MR, Wu X, Wilkison A, Wei Q. Chapter 33: Cancer of the lung: 638-658. Schottenfeld D, Fraumeni JF, Jr. Cancer Epidemiology and Prevention, Third ed. New York: Oxford University Press; 2006.

4. Canadian Cancer Society's Steering Committee on Cancer Statistics. Canadian Cancer Statistics 2011. Toronto: Canadian Cancer Society; 2011.

5. Wu X, Yang H, Lin J. Chapter 3: Lung cancer susceptibility and risk assessment models. 33-60. Roth JA, Cox JD, Hong WK. Lung Cancer. Third ed. Malden, Massachusetts: Blackwell Publishing; 2008.

6. Molina JR, Yang PG, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. Mayo Clinic Proceedings 2008;83(5):584-594.

7. Smith-Warner SA, Ritz J, Hunter DJ, Albanes D, Beeson WL, van den Brandt PA, et al. Dietary fat and risk of lung cancer in a pooled analysis of prospective studies. Cancer Epidemiology, Biomarkers & Prevention 2002;11:987-992.

8. Goodman MT, Hankin JH, Wilkens LR, Kolonel LN. High-Fat Foods and the Risk of Lung Cancer. Epidemiology 1992;3(4):288-299.

9. Alavanja MC, Brown CC, Swanson C, Brownson RC. Saturated Fat Intake and Lung Cancer Risk Among Nonsmoking Women in Missouri. Journal of the National Cancer Institute 1993;85(23):1906-1922.

10. Williams MT, Hord NG. The role of dietary factors in cancer prevention: beyond fruits and vegetables Nutrition in clinical practice 2005;20(4):861-869.

11. Key TJ, Allen NE, Spencer EA, Travis RC. The effect of diet on risk of cancer. The Lancet 2002;360(9):861-867.

12. IARC. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 51. Coffee, tea, mate, methylxanthines and methylglyoxal. Lyon: IARC (International Agency for Research on Cancer); 1991.

13. Al-Wadei HA, Takahashi T, Schuller HM. Caffeine stimulates the proliferation of human lung adenocarcinoma cells and small airway epithelial cells via activation of PKA, CREB and ERK1/2. Oncology reports 2006;15(2):431-436.

14. Lee AH, Liang W, Hirayama F, Binns CW. Association between green tea consumption and lung cancer risk. Journal of preventive medicine and public health 2010;43(4):366-373.

15. Tang NP, Wu YM, Ma J, Wang B, Yu RB. Coffee consumption and risk of lung cancer: A meta-analysis. Lung Cancer 2010;67(1):17-22.

16. National cancer institute. NCI dictionary of cancer terms. 2012. Retrieved from http://www.cancer.gov/cancertopics/types/lung. Assessed on January 2012.

17. Johnson DH, Blot WJ, Carbone DP, Gonzalez A, Hallahan D. Chapter 76: Cancer of the lung: non-small cell lung cancer and small cell lung cancer. Johnson DH, Blot WJ, Carbone DP. Abeloff's Clinical Oncology, fourth ed. Philadelphia: Livingstone Elsevier; 2008.

18. Cooper S, Spiro SG. Small cell lung cancer: Treatment review. Respirology 2006;11(3):241-249.

19. Matukala NT, Parent M, Siemiatycki J, Rousseau M. Socioeconomic Position and Lung Cancer Risk: How Important is the Modeling of Smoking? Epidemiology 2012;23(3):377-85.

20. Berger M, Lung MJ, Brawley OW. Racial Disparities in Lung Cancer. Current problems in cancer 2007;31(3):202-212.

21. Kamp DW. Asbestos-induced lung diseases: an update. Translational Research 2009;153(4):143-152.

22. Goodman M, Morgan RW, Ray R, Malloy CD, Zhao K. Cancer in asbestos-exposed occupational cohorts: a meta-analysis. Cancer Causes & Control 1999;10(5):453-465.

23. Suvatne J, Browning RF. Asbestos and lung cancer. Disease-A-Month 2011;57(1):55-68.

24. Benedetti A, Parent ME, Siemiatycki J. Consumption of alcoholic beverages and risk of lung cancer: results from two case-control studies in Montreal, Canada. Cancer Causes & Control 2006;17(4):469-480.

25. Tea association of Canada. Learn all about tea, 2010. Retrieved from http://www.tea.ca/Page.asp?PageID=1224&SiteNodeID=119&BL_ExpandID=56. Assessed on January 2012.

26. Wickremasinghe RL. Chapter 7: Tea. 229-286. Chichester CO. Advance in food research. Vol. 24. New York: Academic Press; 1978.

27. Balentine DA, Wiseman SA, Bouwens LC. The chemistry of tea flavonoids. . Critical reviews in food Science and nutrition 1997;37(8):693-704.

28. Kim S, David JB. Tea consumption may improve biomarkers of insulin sensitivity and risk factors for diabetes. The Journal of nutrition 2008;138(8):1584S-1592S.

29. Saleem M, Adhami VM, A SI, Mukhtar H. Tea beuverage in chemoprevention of prostate cancer : a mini review. Nutrition and cancer 2003;47(1):13-23.

30. O-Cha.com Japanese Green Tea & Matcha. How green tea processing is accomplished. 2012. Retrieved from http://www.o-cha.com/green-tea-processing.htm. Assessed on January 2012.

31. Singh BN, Shankar S, Srivastava RK. Green tea catechin, epigallocatechin-3-gallate (EGCG): Mechanisms, perspectives and clinical applications. Biochemical pharmacology 2011;82(12):1807-1828.

32. Yang CC, Wang H. Mechanistic issues concerning cancer prevention by tea catechins. Molecular nutrition & food research 2011;55(6):819-850.

33. International Coffee Organization. The story of coffee. 2012. Retrieved from http://www.ico.org/coffee story.asp. Assessed on January 2012.

34. Agriculture and Agri-food Canada. The Canadian coffee industry. 2010. Retrieved from http://www4.agr.gc.ca/AAFC-AAC/display-afficher.do?id=1172237152079&lang=eng#s3. Assessed on January 2012.

35. Coffee Association of Canada. 2010 Canadian coffee drinking study highlight. 2012. Retrieved from http://www.coffeeassoc.com/coffeeincanada.htm. Assessed on January 2012.

36. Coffee Research Institute. Arabica and Robusta coffee plant. 2006. Retrieved from http://www.coffeeresearch.org/agriculture/coffeeplant.htm. Assessed on January 2012.

37. National coffee association USA. Ten step to coffee: From the seed to the cup. 2012. Retrieved from http://www.ncausa.org/i4a/pages/index.cfm?pageid=69. Assessed on January 2012

38. Rajavelu A, Tulyasheva Z, Jaiswal R, Jeltsch A, Kuhnert N. The inhibition of the mammalian DNA methyltransferase 3a (Dnmt3a) by dietary black tea and coffee polyphenols. BMC Biochemistry 2011;12(16).

39. Serafini M, Testa MF. Redox ingredients for oxidative stress prevention: the unexplored potentiality of coffee. Clinics in dermatology 2009;27(7):225-229.

40. Kubik AK, Zatloukal P, Tomasek L, Pauk N, Havel L, Krepela E, et al. Dietary habits and lung cancer risk among non-smoking women. European Journal of Cancer Prevention 2004;13(6):471-480.

41. Kubik A, Zatloukal P, Tomasek L, Kriz J, Petruzelka L, Plesko I. Diet and the risk of lung cancer among women - A hospital-based case-control study. Neoplasma 2001;48(4):262-266.

42. Kubik A, Zatloukal P, Tomasek L, Pauk N, Petruzelka L, Plesko I. Lung cancer risk among nonsmoking women in relation to diet and physical activity. Neoplasma 2004;51(2):136-143.

43. Kubik A, Zatloukal P, Tomasek L, Pauk N, Havel L, Dolezal J, et al. Interactions between smoking and other exposures associated with lung cancer risk in women: diet and physical activity. Neoplasma 2007;54(1):83-88.

44. Kubik A, Zatloukal P, Tomasek L, Dolezal J, Syllabova L, Kara J, et al. A case-control study of lifestyle and lung cancer associations by histological types. Neoplasma 2008;55(3):192-199.

45. Fu YY, Takezaki T, Tajima K. Risk factors of lung cancer—follow-up studies in Nagoya Japan. Zhonghua Liu Xing Bing Xue Za Zhi 1997;18(6):328-330.

46. Zheng W, Doyle TJ, Kushi LH, Sellers TA, Hong CP, Folsom AR. Tea consumption and cancer incidence in a prospective cohort study of postmenopausal women. American Journal of Epidemiology 1996;144(2):175-182.

47. Jacobsen BK, Bjelke E, Kvale G, Heuch I. Coffee drinking, mortality, and cancer incidence: results from a Norwegian prospective study. Journal of the National Cancer Institute 1986;76(5):823-831.

48. Nakachi K, Matsuyama S, Miyake S, Suganuma M, Imai K. Preventive effects of drinking green tea on cancer and cardiovascular disease. Epidemiological evidence for multiple targeting prevention. Biofactors 2000;13(1-4):49-54

49. Zhong LJ, Goldberg MS, Gao YT, Hanley JA, Parent ME, Jin F. A population-based case-control study of lung cancer and green tea consumption among women living in Shanghai, China. Epidemiology 2001;12(6):695-700.

50. Le Marchand L, Murphy SP, Hankin JH, Wilkens LR, Kolonel LN. Intake of Flavonoids and Lung Cancer. Journal of the national cancer institute 2000;92(2):154-160.

51. Bonner MR, Rothman N, Mumford JL, He XZ, Shen M, Welch R, et al. Green tea consumption, genetic susceptibility, PAH-rich smoky coal, and the risk of lung cancer. Mutation Research - Genetic Toxicology & Environmental Mutagenesis 2005;582(1-2):53-60.

52. Khan MMH, Goto R, Kobayashi K, Suzumura S, Nagata Y, Sonoda T. Dietary habits and cancer mortality among middle aged and older Japanese living in hokkaido, Japan by cancer site and sex. Asian Pacific journal of cancer prevention 2004;5(1):58-65.

53. Li Q, Kakizaki M, Kuriyama S, Sone T, Yan H, Nakaya N. Green tea consumption and lung cancer risk: the Ohsaki study. British journal of cancer 2008;99(7):1179-1184.

54. Takezaki T, Hirose K, Inoue M, Hamajima N, Yatabe Y, Mitsudomi T. Dietary factors and lung cancer risk in Japanese: with special reference to fish consumption and adenocarcinomas. British journal of cancer 2001;84(9):1199-1206.

55. Xu ZY, Pan GW, Liu LM, Brown LM, Guan DX, Xiu Q, et al. Cancer risks among iron and steel workers in Anshan, China. 1. Proportional mortality ratio analysis. American Journal of Industrial Medicine 1996;30(1):1-6.

56. Ohno Y, Wakai K, Genka K, Ohmine K, Kawamura T, Tamakoshi A, et al. Tea consumption and lung cancer risk: a case-control study in Okinawa, Japan. Japanese journal of cancer research 1995;86(11):1027-1034.

57. Hirvonen T, Virtamo J, Korhonen P, Albanes D, Pietinen P. Flavonol and flavone intake and the risk of cancer in male smokers (Finland). Cancer causes and control 2001;12(9):789-796.

58. Hu JF, Mao Y, Dryer D, White K. Risk factors for lung cancer among Canadian women who have never smoked. Cancer Detection & Prevention 2002;26(2):129-138.

59. Baker JA, McCann SE, Reid ME, Nowell S, Beehler GP, Moysich KB. Associations between black tea and coffee consumption, and risk of lung cancer among current and former smokers. Nutrition & Cancer - an International Journal 2005;52(1):15-21.

60. Goldbohm RA, Hertog MG, Brant HA, Poppel GV, Brandt PVD. Consumption of Black Tea and Cancer Risk: a Prospective Cohort Study. Journal of the National Cancer institute 1996;88(2):93-100.

61. Cui Y, Morgenstern H, Greenland S, Tashkin DP, Mao JT, Cai L, et al. Dietary flavonoid intake and lung cancer - A population-based case-control study. Cancer 2008;112(10):2241-2248.

62. Mettlin C. Milk drinking, other beverage habits, and lung cancer risk. International journal of cancer 1989;43(4):608-612.

63. Sankaranarayanan R, Varghese C, Duffy SW, Padmakumary G, Day NE, Nair MK. A case-control study of diet and lung cancer in Kerala, south India. International journal of cancer 1994;58(5):644-649.

64. Nyberg F, Agrenius V, Svartengren K, Svensson C, Pershagen G. Dietary factors and risk of lung cancer in never-smokers. International Journal of Cancer 1998;78(4):430-436.

65. Kinlen LJ, Willows AN, Goldblatt P, Yudkin J. Tea consumption and cancer. . British Journal of Cancer 1988;58(3):397-401.

66. Axelsson G, Liljeqvist T, Andersson L, Bergman B, Rylander R. Dietary factors and lung cancer among men in west Sweden. International Journal of Epidemiology 1996;25(1):32-39.

67. Mendilaharsu M, De Stefani E, Deneo-Pellegrini H, Carzoglio JC, Ronco a. Consumption of tea and coffee and the risk of lung cancer in cigarette-smoking men: a case-control study in Uruguay. Lung cancer 1998;19(2):101-107.

68. Lin Y, Cai L. Environmental and Dietary Factors and Lung Cancer Risk Among Chinese Women: A Case-Control Study in Southeast China. Nutrition and cancer 2012;64(4):508-514.

69. Stensvold I, Jacobsen BK. Coffee and cancer - a prospective study of 43,000 Norwegian men and women. Cancer Causes & Control 1994;5(5):401-408.

70. Chiu YL, Wang XR, Qiu H, Yu ITS. Risk factors for lung cancer: a case-control study in Hong Kong women. Cancer Causes & Control 2010;21(5):777-785.

71. Chow WH, Schuman LM, McLaughlin JK, Bjelke E, Gridley G, Wacholder S, et al. A Cohort Study of Tobacco Use, Diet, Occupation, and Lung Cancer Mortality. Cancer Causes & Control 1992;3(3):247-254.

72. Hartman TJ, Tangrea JA, Pietinen P, Malila N, Virtanen M, Taylor PR, et al. Tea and coffee consumption and risk of colon and rectal cancer in middle-aged finnish men. Nutrition & Cancer - an International Journal 1998;31(1):41-48.

73. Weisburger JH. Tea and health: a historical perspective. Cancer letters 1997;114(1-2):315-317.

74. Mukhtar H, Ahmad N. Tea polyphenols: Prevention of cancer and optimizing health. American Journal of Clinical Nutrition 2000;71(6 suppl):1698S-1702S.

75. India Trade Promotion Organisation. Tea market in Sweden. 2009. Retrieved from http://www.tradeportalofindia.com/usrdata/webadmin/Section3.9/HS09_Tea_0302.htm. Assessed on January 2012.

76. Agriculture and Agri-food Canada. The Canadian Tea industry. 2010. Retrieved from http://www4.agr.gc.ca/AAFC-AAC/display-afficher.do?id=1298047470064&lang=eng. Assessed on January 2012.

77. Hakim IA, Harris RB, Weisgerber UM. Tea Intake and Squamous Cell Carcinoma of the Skin: Influence of Type of Tea Beverages. Cancer Epidemiology, Biomarkers & Prevention 2000;9(7):727-731.

78. International Trade Centre. The coffee guide: trade practices or relevance to exporters in coffee-producing countries. 2010. Retrieved from http://www. thecoffeeguide. org /coffee-guide/the-markets-for-coffee/ Decaffeinated-coffee/. Assessed on January 2012.

79. Darmon N, Drewnowski A. Does social class predict diet quality? The American journal of clinical nutrition 2008;87(5):1107-17.

80. Maynard M, Gunnell D, Ness AR, Abraham L, Bates CJ, Blane D. Related Behaviours What influences diet in early old age? Prospective and cross-sectional analyses of the Boyd Orr cohort. European Journal of public health 2006;16(3):316-324.

81. Smith AM, Baghurst KI. Public health implications of dietary differences between social status and occupational category groups. Journal of Epidemiology and Community Health 1992;46(2):409-416.

82. Statistic Canada. Census tract. 2009. Retreived from http://www12.statcan.ca/ census-recensement/2006/ref/dict/geo013a-eng.cfm. Assessed on November 2012.

83. Nellis MJ, Emurian HH, Brady JV, Ray RL. Behavior analysis of cigarette smoking. The Pavlovian journal of biological science 1982;17(3):140-149.

84. Inoue M, Kurahashi N, Iwasaki M, Shimazu T, Tanaka Y, Mizokami M, et al. Effect of coffee and green tea consumption on the risk of liver cancer: cohort analysis by hepatitis virus infection status. Cancer Epidemiology, Biomarkers & Prevention 2009;18(6):1746-1753.

85. Rodenburg EM, Eijgelsheim M, Geleijnse JM, Amin N, Duijn CMV, Hofman A. CYP1A2 and coffee intake and the modifying effect of sex, age , and smoking. The American journal of clinical nutrition 2012;96(1):182-187.

86. Leffondré K, Abrahamowicz M, Xiao Y, Siemiatycki J. Modelling smoking history using a comprehensive smoking index: application to lung cancer. Statistics in Medicine 2006;25(24):4132-4146.

87. Dietrich T, Hoffmann K. A comprehensive index in the modeling of smoking history in periodontal research. Journal of dental research 2004;83(12):859-863.

88. Smith-Warner SA, Spiegelman D, Yaun SS, Albanes D, Beeson WL, van den Brandt PA, et al. Fruits, vegetables and lung cancer: A pooled analysis of cohort studies. International Journal of Cancer 2003;107(6):1001-1011.

89. Bandera EV, Freudenheim JL, Vena JE. Alcohol consumption and lung cancer: A review of the epidemiologic evidence. Cancer Epidemiology, Biomarkers & Prevention 2001;10(8):813-821.

90. Hamer M, Witte DR, Mosdøl A, Marmot MG, Brunner EJ. Prospective study of coffee and tea consumption in relation to risk of type 2 diabetes mellitus among men and women : The Whitehall II study. British Journal of Nutrition 2008;100(5):1046-1053.

91. Willett WC, Howe R, Kushi LH. Adjustment for total energy intake in epidemiologic studies. The American journal of clinical nutrition. 1997;65(4 suppl):1220S-1228S.

92. Lavoué J. The Montréal Lung study exposure data Description and creation of the Job and Subjects exposure files. 2011.

93. Baneshi MR, Talei AR. Does the Missing Data Imputation Method Affect the Composition and Performance of Prognostic Models ? . Iranian Red Crescent medical journal 2012;14(1):31-36.

94. Kristman VL, Manno M, Co P. Methods to account for attrition in longitudinal data: Do they work? A simulation study. European Journal of Epidemiology 2005;20(8):657-662.

95. Heijden GJVD, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research : A clinical example. Journal of Clinical Epidemiology 2006;59(10):1102-1109.

96. Kleinbaum DG, Klein M. Chapter 1: Introduction to logistic regression: 32-41. Logistic regression a self-learning text, third ed. New York: Springler; 2010.

97. Zhang M, Binns CW, Lee AH. Tea Consumption and Ovarian Cancer Risk: A Case-Control Study in China. Cancer epidemiology, biomarkers and prevention 2002;11(8):713-718.

98. Rees JR, Stukel TA, Perry AE, Zens MS, Spencer SK, Karagas MR. Tea consumption and basal cell and squamous cell skin cancer: Results of a case-control study. Journal of the American Academy of Dermatology 2007;56(5):781-785.

99. Sala M, Cordier S, Chang-Claude J, Donato F, Escolar-Pujolar A, Fernandez F, et al. Coffee consumption and bladder cancer in nonsmokers: a pooled analysis of case-control studies in European countries. Cancer Causes & Control 2000;11(10):925-931.

100. Schwarz B, Bischof HP, Kunze M. Coffee, tea, and lifestyle. Preventive medicine. 1994;23(3):377-384.

101. Freedman ND, Park Y, Abnet CC, Hollenbeck AR, Sinha R. Association of Coffee Drinking with Total and Cause-Specific Mortality. The New England journal of medicine 2012;366(20):1891-1904.

Appendix

Appendix 1: Summaries of the studies identified for the literature review

Table 1: Summaries of the studies identified for the literature review on the association between black tea and lung cancer

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Hirvonen et al (2001)(57)	Cohort	Males	Smokers	791	Age, supplementation group, years of smoking, daily amount of cigarettes	≥ 1 cup/day vs < 1 cup/day	0.66	0.54-0.82	-
Kinlen et al (1988) (65)	Cohort	Males	Both	718	Age	0 to 3 cups a day	0.61[a]	-	<0.0001
						4 to 6 cups a day	0.80[a]	-	
						7 to 9 cups a day	1.13[a]	-	
						≥ 10 cups a day	1.41[a]	-	
Khan et al (2004) (52)	Cohort	Males	Both	41	Age, smoking	≥several times a week vs ≤ several times a month	0.6	0.1 - 2.5	-
		Females	Both	10	Age, health status, health education, health screening, smoking	≥several times a week vs ≤ several times a month	2.1	0.3 - 17.5	
Goldbohm et al (1996) (60)	Cohort	Both	Both	676	Age, sex, education, smoking status, pack-years of cigarettes smoked, family history of lung cancer, intake of beta-carotene and vitamin C	1 cup/day vs non-consumer	0.93	0.64 - 1.36	0.910
						2 cups/day vs non-consumer	1.01	0.71 - 1.41	
						3 cups/day vs non-consumer	0.84	0.56 - 1.26	
						4 cups/day vs non-consumer	0.9	0.62 - 1.31	
						5 cups/day vs non-consumer	1.07	0.73 - 1.57	
Hu et al (2002) (58)	CC	Females	N-smokers	153	10-year age group, province, education, social class	1 to 7 cups/week vs 0 cup/week	0.6	0.3 - 0.9	0.0008
						>7 cups/week vs 0 cup/week	0.4	0.2 - 0.7	

CC: Case-control

OR/RR: Odd ratio/risk ratio

95%CI: 95% confidence interval

P-trend: P-value for the trend test

a: Standardized Mortality Ratio, calculated using indirect standardization.

Table 1: Summaries of the studies identified for the literature review on the association between black tea and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Cui et al (2008) (61)	CC	Both	Both	558	Age, sex, race-ethnicity, years of schooling, smoking status, pack-years of tobacco smoking, daily energy intake	>0 to 1 cup/day vs 0 cup/day	0.64	0.49 - 0.85	-
						>1 cup/day vs 0 cup/day	0.42	0.24 - 0.73	
			Smokers	462		>0 to 1 cup/day vs 0 cup/day	0.61	0.44 - 0.85	
						>1 cup/day vs 0 cup/day	0.37	0.19 - 0.72	
			N-smokers	96	Age, sex, race-ethnicity, years of schooling, daily energy intake	>0 to 1 cup/day vs 0 cup/day	0.83	0.49 - 1.40	
						>1 cup/day vs 0 cup/day	0.52	0.18 - 1.50	
Mendilaharsu et al (1998) (67)	CC	Males	Smokers	427	Age, residence, urban/rural status, pack-years tobacco smoking, total energy intake, dairy foods, desserts, all vegetables and fruits, mate intake, caffeine index, coffee	ever drinkers vs non-drinkers	0.73	0.51 - 1.04	0.028
						1 cup/week vs non-drinkers	0.69	0.44 - 1.11	
						2 - 3 cups/week vs non-drinkers	0.94	0.59 - 1.51	
						1 cup/day vs non-drinkers	0.51	0.28 - 0.95	
						≥ 2 cups/day vs non-drinkers	0.34	0.14 - 0.84	
			< 44 pack-years	126		ever drinkers vs non-drinkers	0.48	0.3 - 0.8	0.002
						1 cup/week vs non-drinkers	0.53	0.3 - 1.1	
						2 - 3 cups/week vs non-drinkers	0.74	0.4 - 1.5	
						≥ 1 cup/day vs non-drinkers	0.3	0.1 - 0.6	
			≥ 45 pack-years	301		ever drinkers vs non-drinkers	0.98	0.6 - 1.5	0.950
						1 cup/week vs non-drinkers	0.77	0.4 - 1.4	
						2 - 3 cups/week vs non-drinkers	1.45	0.8 - 2.7	
						≥ 1 cup/day vs non-drinkers	0.82	0.4 - 1.5	
Sankaranarayanan et al (1994) (63)	CC	Males	Both	251	Age, education, religion, smoking	Ever vs never	3.93	2.55 - 5.92	<0.001

Table 1: Summaries of the studies identified for the literature review on the association between black tea and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Kubik et al (2008) (44)	CC	Females	Smokers	383	Age , residence, education, pack-years of smoking	at least once per week vs never	1.24	0.93 - 1.66	
			N-smokers	163		at least once per week vs never	0.69	0.49 - 0.98	
			Both	207[b]		at least once per week vs never	0.98	0.72 - 1.33	-
			Both	149[c]		at least once per week vs never	1.32	0.91 - 1.91	
			Both	132[d]		at least once per week vs never	0.94	0.63 - 1.41	
		Males	Smokers	403		at least once per week vs never	1.00	0.70 - 1.41	
			N-smokers	66		at least once per week vs never	1.51	0.77 - 2.94	
			Both	101[b]		at least once per week vs never	1.17	0.72 - 1.90	-
			Both	249[c]		at least once per week vs never	1.27	0.87 - 1.85	
			Both	81[d]		at least once per week vs never	0.93	0.56 - 1.57	
Axelsson et al (1996) (66)	CC	Males	Both	308	Age, cigarettes/day, number of years smoked, marital status, socioeconomic job classification, vegetable class, other fruit or berries	once or twice a week vs twice a month or less	0.92	0.57 - 1.50	
						daily or almost daily vs twice a month or less	1.19	0.78 - 1.83	-
						several times/day vs twice a month or less	0.74	0.33 - 1.64	
Le Marchand et al (2000) (50)	CC	Both	Both	582	Age, sex and ethnicity, smoking status, duration of smoking, cigarettes per day, intake of B-carotene, intake of saturated fat	Quartile 2 vs quartile 1	1.5	0.8 - 2.6	
						Quartile 3 vs quartile 1	1.1	0.7 - 1.8	0.83
						Quartile 4 vs quartile 1	1.1	0.7 - 1.8	
Mettlin et al (1989) (62)	CC	Both	Both	569	Age strata and residence, Sex, smoking history, beta-carotene intake index, education level	<1 cup/day vs never	0.86	0.63 - 1.15	
						2-3 cups/day vs never	0.86	0.56 - 1.33	-
						≥ 4 cups/day vs never	1.09	0.62 - 1.93	

b: Only include adenocarcinoma

c: Only include squamous cell carcinoma

d: Only include small cell carcinoma

Table 1: Summaries of the studies identified for the literature review on the association between black tea and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Baker et al (2005) (59)	CC	Both	Both	993	Age, sex, smoking status, known occupational exposure to asbestos, number of cigarettes smoked per day, interaction between asbestos and cigarettes	<1 cup a day vs none	0.87	0.66 - 1.16	0.93
						1 cup a day vs none	0.97	0.66 - 1.43	
						≥ 2 cups a day vs none	0.90	0.66 - 1.24	
				120[b]		<1 cup a day vs none	0.85	0.56 - 1.58	-
						1 cup a day vs none	0.81	0.34 - 1.96	
						≥ 2 cups a day vs none	0.71	0.35 - 1.46	
				170[e]		<1 cup a day vs none	0.92	0.55 - 1.55	-
						1 cup a day vs none	0.88	0.41 - 1.89	
						≥ 2 cups a day vs none	0.82	0.47 - 1.45	
				186[d]		<1 cup a day vs none	0.94	0.57 - 1.57	-
						1 cup a day vs none	1.06	0.52 - 2.15	
						≥ 2 cups a day vs none	0.65	0.35 - 1.19	
				366[c]		<1 cup a day vs none	0.81	0.55 - 1.19	-
						1 cup a day vs none	0.75	0.44 - 1.29	
						≥ 2 cups a day vs none	0.88	0.58 - 1.34	
Nyberg et al (1998) (64)	CC	Both	N-smokers	124	Age, gender, catchment area, occasional smoking, degree of urban residences, years of exposure to environmental tobacco smoke, carrot consumption, other fruits consumption	1-4 times a week vs less than weekly	1.04	0.56 - 1.95	0.37
						daily or almost daily vs less than weekly	1.27	0.73 - 2.21	
Ohno et al[f] (1995) (56)	CC	Males	Both	39	Age, residence, education, smoking, family history of lung cancer, medical history of lung disease, intake frequency of green-yellow vegetables	Ever vs never	1.15	0.70 - 1.91	-
		Females		5		Ever vs never	0.10	0.02 - 0.46	

e: Only include large cell carcinoma

f: Only include subjects consuming green tea

Table 2: Summaries of the studies identified for the literature review on the association between coffee and lung cancer

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Chow et al (1992) (71)	Cohort	Males	Both	219	Age, industry/occupation, smoking	3-4 cups/day vs <3 cups/day	2.1	1.4 - 3.2	-
						5-6 cups/day vs <3 cups/day	2.1	1.3 - 3.3	
						>6 cups/day vs <3 cups/day	2.4	1.4 - 4.2	
Khan et al (2004) (52)	Cohort	Males	Both	41	Age, smoking	≥several times a week vs ≤ several times a month	0.7	0.4 - 1.4	-
		Females		10	Age, health status, health education, health screening, smoking	≥several times a week vs ≤ several times a month	2.1	0.5 - 8.0	
Stensvold et al (1994) (69)	Cohort	Males	Both	93	Age, cigarettes per day, county of residence	5-6 cups/day vs ≤4 cups/day	1.4	p > 0.05[a]	Significant
						≥ 7 cups/day vs ≤4 cups/day	2.4	p < 0.01[a]	
		Females		32		5-6 cups/day vs ≤4 cups/day	1.9	p > 0.05[a]	Significant
						≥ 7 cups/day vs ≤4 cups/day	2.0	p > 0.05[a]	
Chiu et al (2010) (70)	CC	Females	Both	279	Age, employment, years of education, total dish-years, smoking, family cancer history, radon index	1-10 cups per day for 1 year vs none	0.41	0.21 - 0.78	0.005
						>10 cups per day for 1 year vs none	1.30	0.70 - 2.42	
Axelsson et al (1996) (66)	CC	Males	Both	308	Age, cigarettes/day, number of years smoked, marital status, socioeconomic job classification, vegetable class, other fruits or berries	daily or almost daily vs twice a week or less	0.94	0.38 - 2.29	-
						7-25 times/week vs twice a month or less	1.16	0.53 - 2.52	
						>25 times/week vs twice a month or less	1.60	0.72 - 3.54	
Sankaranarayanan et al (1994) (63)	CC	Males	Both	251	Age, education, religion, smoking	Ever vs never	0.69	0.48 - 1.01	0.05

CC: Case-control

OR/RR: Odd ratio/risk ratio

95%CI: 95% confidence interval

P-trend: P-value for the trend test

a: 95% confidence intervals not given

Table 2: Summaries of the studies identified for the literature review on the association between coffee and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Kubik et al (2008) (44)	CC	Females	Smokers	383	Age , residence, education, pack-years of smoking	at least once per week vs never	0.76	0.48 - 1.20	-
			N-smokers	163		at least once per week vs never	0.86	0.59 – 1.26	
			Both	207[b]		at least once per week vs never	0.93	0.62 – 1.38	
			Both	149[c]		at least once per week vs never	0.80	0.49 – 1.30	
			Both	132[d]		at least once per week vs never	0.90	0.52 – 1.55	
		Males	Smokers	403		at least once per week vs never	1.07	0.61 – 1.86	
			N-smokers	66		at least once per week vs never	0.91	0.43 – 1.92	
			Both	101[b]		at least once per week vs never	1.58	0.74 – 3.36	
			Both	249[c]		at least once per week vs never	0.98	0.57 – 1.66	
			Both	81[d]		at least once per week vs never	0.93	0.45 – 1.92	
Mendilaharsu et al (1998) (67)	CC	Males	Smokers	427	Age, residence, urban/rural status, pack-years tobacco smoking, total energy intake, dairy foods, desserts, all vegetables and fruits, mate intake, caffeine index, coffee	ever drinkers vs non-drinkers	1.11	0.66 -1.85	0.72
						1 cup/week vs non-drinkers	1.36	0.71 - 2.62	
						2 - 3 cups/week vs non-drinkers	0.87	0.45 - 1.69	
						1 cup/day vs non-drinkers	1.11	0.49 - 2.52	
						≥ 2 cups/day vs non-drinkers	1.06	0.40 - 2.81	
Nyberg et al (1998) (64)	CC	Both	N-smokers	124	Age, gender, catchment area, occasional smoking, degree of urban residences, years of exposure to environmental tobacco smoke, carrot consumption, other fruits consumption	daily or almost daily vs less than daily	0.57	0.27 - 1.22	0.33
						≥ 3 cups/day vs less than daily	0.50	0.24 - 1.06	

b: Only include adenocarcinoma

c: Only include squamous cell carcinoma

d: Only include small cell carcinoma

Table 2: Summaries of the studies identified for the literature review on the association between coffee and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Hu et al (2002) (58)	CC	Females	N-smokers	155	Age, province, education, social class	2-7 cups/week vs ≤1 cup/week	0.9	0.5 - 1.6	
						8 - 17.5 cups/week vs ≤1 cup/week	0.9	0.5 - 1.6	0.67
						>17.5 cups/week vs ≤1 cup/week	0.8	0.4 - 1.8	
Mettlin et al (1989) (62)	CC	Both	Both	569	Age , residence, Sex, smoking history, beta-carotene intake index, education level	<1 cup/day vs never	1.01	0.67 - 1.51	
Takezaki et al (2001) (54)	CC	Males	Both	367[b]	Age, season and year of visit, occupation, prior lung disease, smoking, consumption of green vegetables and meat	1 cup/day vs < 1cup/day	0.85	0.61 - 1.19	
						2 cups/day vs < 1cup/day	0.87	0.60 - 1.25	0.654
						≥3 cups/day vs < 1cup/day	1.18	0.80 - 1.74	
				381[e]		1 cup/day vs < 1cup/day	0.98	0.70 - 1.37	
						2 cups/day vs < 1cup/day	1.15	0.80 - 1.64	0.027
						≥3 cups/day vs < 1cup/day	1.61	1.09 - 2.39	
		Females		240[b]		1 cup/day vs < 1cup/day	0.76	0.51 - 1.13	
						2 cups/day vs < 1cup/day	0.82	0.49 - 1.35	0.823
						≥3 cups/day vs < 1cup/day	1.28	0.65 - 2.54	
				57[e]		1 cup/day vs < 1cup/day	0.96	0.43 - 2.18	
						2 cups/day vs < 1cup/day	0.61	0.21 - 1.78	0.142
						≥3 cups/day vs < 1cup/day	0.28	0.05 - 1.58	

e: Only include squamous cell carcinoma and small cell carcinomas

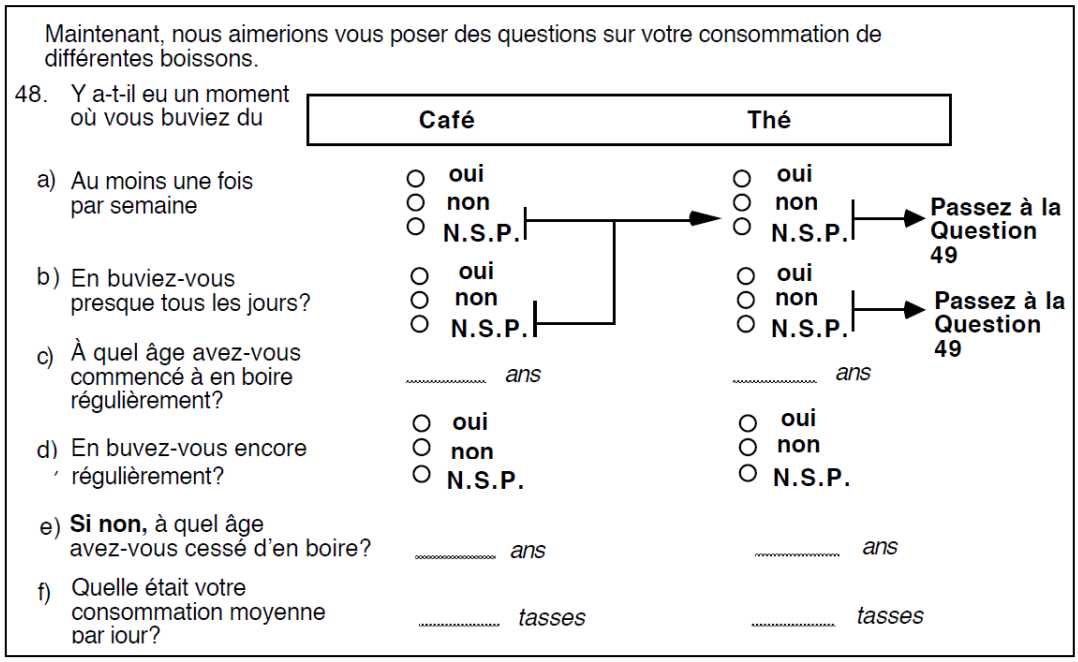
Table 2: Summaries of the studies identified for the literature review on the association between coffee and lung cancer (continued)

	Design	Sex	smoking status	Cases (n)	Adjusted for	Exposure	OR/RR	95%CI	p-trend
Baker et al (2005) (59)	CC	Both	Both	972	Age, sex, smoking status, known occupational exposure to other type of dust, known occupational exposure to smoke, number of cigarettes smoked per day, interaction between smoke exposure and cigarettes	≤1 cup/day vs none	1.03	0.73 - 1.45	0.01
						2-3 cups/day vs none	1.34	0.99 - 1.82	
						≥ 4 cups/day vs none	1.51	1.11 - 2.05	
				118[b]		≤1 cup/day vs none	0.90	0.45 - 1.81	-
						2-3 cups/day vs none	0.60	0.30 - 1.21	
						≥ 4 cups/day vs none	1.66	0.90 - 3.04	
				164[f]		≤1 cup/day vs none	1.12	0.55 - 2.26	-
						2-3 cups/day vs none	1.44	0.78 - 2.66	
						≥ 4 cups/day vs none	1.82	1.00 - 3.29	
				181[d]		≤1 cup/day vs none	0.67	0.33 - 1.37	-
						2-3 cups/day vs none	1.51	0.88 - 2.60	
						≥ 4 cups/day vs none	1.48	0.86 - 2.53	
				360[c]		≤1 cup/day vs none	1.12	0.70 - 1.80	-
						2-3 cups/day vs none	1.28	0.84 - 1.97	
						≥ 4 cups/day vs none	1.61	1.05 - 2.47	

f: Only include large cell carcinoma

Appendix 2: Questions on tea and coffee from the interview

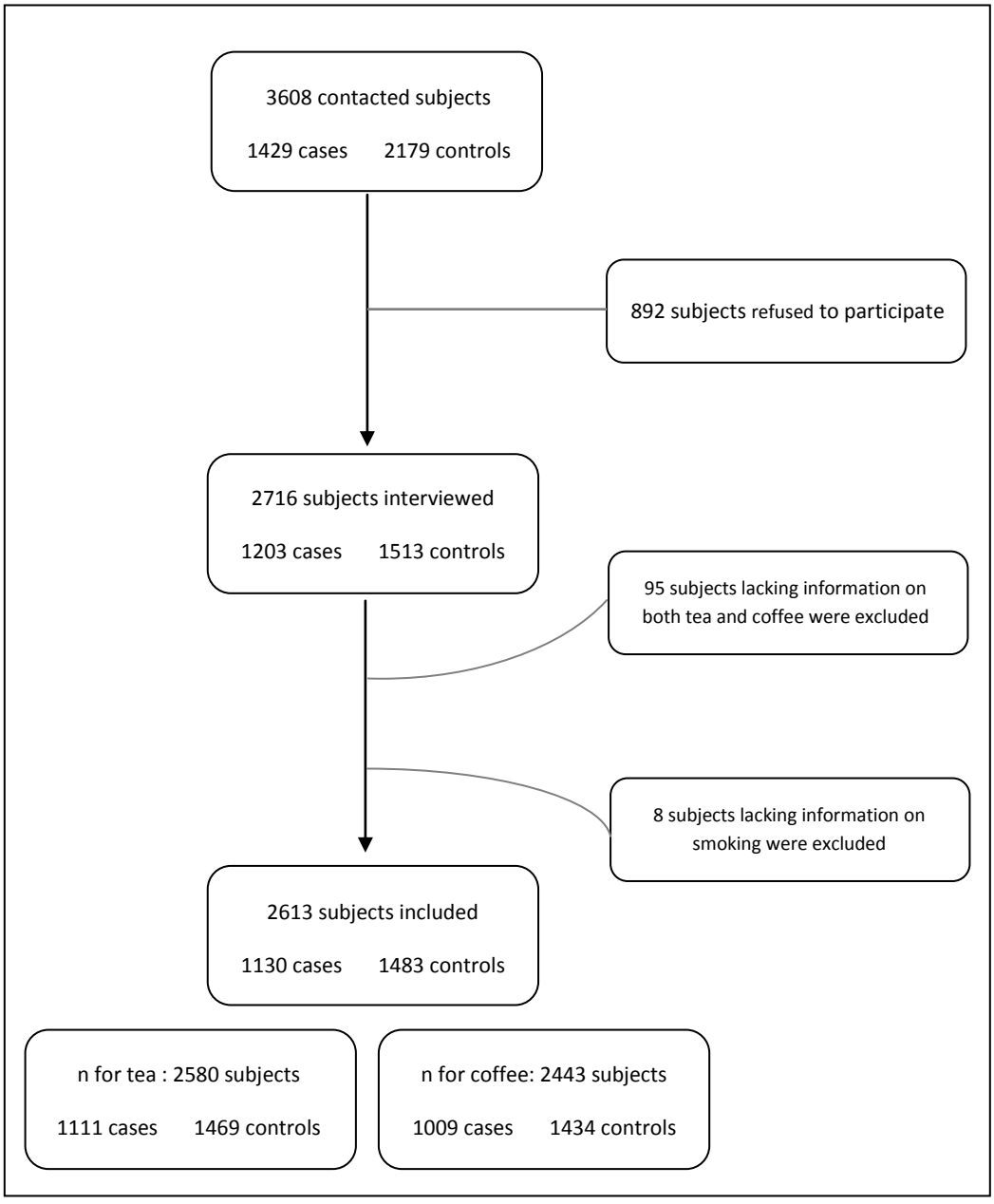
Figure 1: Questionnaire on the consumption of tea and coffee



*N.S.P = ne sais pas/ don't know

Appendix 3: Flowchart of the selection process

Figure 2: Flowchart of the selection process



Appendix 4: Summary of the model reduction procedure

Table 3: Summary of the model reduction procedure for black tea

Tested variable	Categories	OR		% change in OR
		With variable	Without variable	
Spirit intake	Occasional drinkers	0.810	0.811	0.123
	Regular drinkers	1.088	1.089	0.092
Exposure to asbestos	Occasional drinkers	0.811	0.811	0.000
	Regular drinkers	1.089	1.091	0.184
Beer intake	Occasional drinkers	0.811	0.801	-1.233
	Regular drinkers	1.091	1.075	-1.467
Vegetable intake	Occasional drinkers	0.801	0.792	-1.124
	Regular drinkers	1.075	1.058	-1.581
Exposure to silica	Occasional drinkers	0.792	0.810	2.273
	Regular drinkers	1.058	1.071	1.229
Fruit intake	Occasional drinkers	0.810	0.790	-2.469
	Regular drinkers	1.071	1.074	0.280
Wine intake	Occasional drinkers	0.790	0.716	-9.367
	Regular drinkers	1.074	1.079	0.466

*Starting model included: tea frequency of consumption, age, sex, ethnicity, median of family income (from census tract),total years of education, respondent proxy status, smoking (CSI), vegetable intake, fruit intake, beer intake, spirit intake, wine intake, energy intake, exposure to silica, and exposure to asbestos

Table 4: Summary of the model reduction procedure for coffee

Tested variable	Categories	OR		% change in OR
		With variable	Without variable	
Vegetable intake	Occasional drinkers	0.310	0.308	-0.645
	Regular drinkers	0.783	0.778	-0.639
Exposure to asbestos	Occasional drinkers	0.308	0.311	0.974
	Regular drinkers	0.778	0.778	0.000
Spirit intake	Occasional drinkers	0.311	0.307	-1.286
	Regular drinkers	0.778	0.776	-0.257
Beer intake	Occasional drinkers	0.307	0.305	-0.651
	Regular drinkers	0.776	0.765	-1.418
Fruit intake	Occasional drinkers	0.305	0.306	0.328
	Regular drinkers	0.765	0.782	2.222
Exposure to silica	Occasional drinkers	0.306	0.315	2.941
	Regular drinkers	0.782	0.785	0.384
Wine intake	Occasional drinkers	0.315	0.252	-20.000
	Regular drinkers	0.785	0.657	-16.306

*Starting model included: coffee frequency of consumption, age, sex, ethnicity, median of family income (from census tract),total years of education, respondent proxy status, smoking (CSI), vegetable intake, fruit intake, beer intake, spirit intake, wine intake, energy intake, exposure to silica, and exposure to asbestos